THE AMERICAN MATHEMATICAL MONTHLY

AN OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA
(INCORPORATED)

DEVOTED TO THE INTERESTS OF COLLEGIATE MATHEMATICS

THE AMERICAN MATHEMATICAL MONTHLY, FOUNDED IN 1894 BY BENJAMIN F. FINKEL, WAS PUBLISHED BY HIM UNTIL 1913. FROM 1913 TO 1916 IT WAS OWNED AND PUBLISHED BY REPRESENTATIVES OF FOURTEEN UNIVERSITIES AND COLLEGES IN THE MIDDLE WEST

VOLUME 84
1977

PUBLISHED BY THE ASSOCIATION
MENASHA, WIS., AND WASHINGTON, D.C.

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

AN OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 84 | NUMBER 1

CODEN: AMMYAE

CONTENTS

(*Continued on inside cover*)

JANUARY **1977**

NOTICE TO AUTHORS

Research papers per se are unsuitable; see statement of policy (Vol. 81, p. 1).

Please follow the format in current issues of the Monthly. Manuscripts must be legibly typewritten or reproduced from typewritten copy, double spaced with wide margins. Three copies should be submitted to the appropriate editor and one kept by the author as protection against loss. The author's full address *must* appear at the end of the manuscript.

Backlog: Main Articles 18 months, Math. Notes 20 months, Research Problems 16 months, Classroom Notes 16 months, Math. Education 18 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to R. P. Boas, Department of Mathematics, Northwestern University, Evanston, IL 60201; NOTES, etc.: to the corresponding Associate Editor; REPRINT PERMISSION: to Leonard Gillman, Mathematical Association of America, University of Texas, Austin, Texas 78712 (see also the copyright notice below); ADVERTISING CORRESPONDENCE: to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo, Buffalo, N.Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. Willcox, Mathematical Association of America, 1225 Connecticut Ave., N. W., Washington, D.C. 20036; BACK ISSUES: Contact P. and H. Bliss Co., Middletown, CT 06457.

Annual dues for members of the Association, including a subscription to the American Mathematical Monthly, are \$18.00 for each of the first two years of membership and \$22.00 thereafter. Student Membership is available with annual dues of \$12.00. For nonmembers the subscription price is \$25.00.

Published by the Association at Washington, D.C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D.C., and additional mailing offices.

PRINTED IN THE UNITED STATES OF AMERICA

THE LANDAU PROBLEM FOR MOTIONS IN A RING AND IN BOUNDED CONTINUA

I. J. SCHOENBERG

Introduction. The present paper should be accessible to a student who has had a course in introductory Calculus. It deals with Landau's Theorem 1, of Section 1 (see [**5**, Section 4]), and generalizes it to complex valued functions $f(t)$, $(t \in R)$, with values in a ring $K_r = \{z\,;\, r \leqq |z| \leqq 1\}$, where r, $0 \leqq r < 1$, is preassigned. The parabolic motion of a particle in a constant field of force plays a major role.

Similar generalizations of Kolmogorov's theorem (see [**3**], [**5**], and Cavaretta's remarkable paper [**1**]) to motions in a ring, encounter serious obstacles. Even the case of the ring K_r remains unresolved as soon as $n \geqq 3$. The approximate differentiation formulae so successfully used to establish Kolmogorov's theorem (see [**5**] and especially [**2**]) no longer seem effective for the case of a ring, even not when $n = 2$. Presumably, the author's so-called exponential Euler splines (see [**6**, Lecture 3]) are extremizing functions in K_r for a certain interval of values of r, $0 \leqq r \leqq r_{n,k}$, $(0 < r_{n,k} < 1)$, and a precise conjecture might be stated on a future occasion.

A longer version of the present paper appeared under the same title as a Mathematics Research Center T.S.Rep. #1563, in October 1975. The first seven sections of both versions are identical, but the present Section 8 was much expanded in the MRC report. The author is indebted to the referee, Professor A. Cavaretta, for suggesting that the material beyond Section 7 be much reduced. The result is a paper that is better suited for its purpose.

1. Landau's original theorem. Let $f(t)$ be real-valued and defined on the real axis $R = (-\infty, \infty)$. We also assume that $f(t)$ is continuously differentiable, hence $f'(t) \in C(R)$, and that $f'(t)$ is piecewise continuously differentiable. We also assume that the supremum norm

$$\|f''\| = \sup |f''(t)|, \qquad (t \in R)$$

is finite. In 1913 E. Landau [**4**] established a result that is equivalent to the following

THEOREM 1 (E. Landau). *If $f(t)$ satisfies the two conditions*

$$\|f\| \leqq 1, \qquad \|f''\| \leqq 8, \tag{1.1}$$

then

$$\|f'\| \leqq 4, \tag{1.2}$$

and 4 is here the best constant.

Observe that Theorem 1 generalizes itself for the case when $f(t)$ is *complex-valued.* This is readily seen if we think of $f(t)$ as a motion of a point in the complex plane and we reason geometrically in kinematic terms. For if f: $R \to \mathbb{C}$, then $f(t)$ represents a motion that is restricted to the unit circle $U = \{|z| \leqq 1\}$ for all t, and having an acceleration $f''(t)$ such that $|f''(t)| \leqq 8$ for all t. If L is any straight line in the plane, and $\tilde{f}(t)$ is the orthogonal projection of $f(t)$ on L, then $\tilde{f}(t)$ describes a motion on L subject to the two restrictions:

1. $\tilde{f}(t)$ is confined to an interval I of length 2 (I is the projection of U on L),
2. $|\tilde{f}''(t)| \leqq 8$ for all t. By Theorem 1 we conclude that

$$|\tilde{f}'(t)| \leqq 4 \quad \text{for all } t. \tag{1.3}$$

It follows that $|f'(t)| \leqq 4$ for all t, for if $|f'(t_0)| > 4$, we would obtain a contradiction to (1.3) on choosing for L the line carrying the vector $f'(t_0)$.

2. A generalization of Theorem 1. We begin by observing that the function $x/(1+x^2)$ is increasing in the interval $0 \leqq x \leqq 1$. Setting $x = \cos(u/2)$, it follows that the function

$$g(u) = \frac{4\cos(u/2)}{2+2\cos^2(u/2)} = \frac{4\cos(u/2)}{3+\cos u}, \quad (0 \leqq u \leqq \pi), \tag{2.1}$$

decreases from $g(0)=1$ to $g(\pi)=0$ as u increases from 0 to π. If r is prescribed such that

$$0 < r < 1, \tag{2.2}$$

it follows that there is a unique value u such that

$$r = \frac{4\cos(u/2)}{3+\cos u}, \quad (0 < u < \pi). \tag{2.3}$$

Our generalization of Theorem 1 is as follows:

THEOREM 2. *If $f(t)$ is complex-valued and such that*

$$r \leqq |f(t)| \leqq 1 \quad \textit{for all } t, \quad \|f''\| \leqq \frac{16\sin^2(u/2)}{3+\cos u}, \tag{2.4}$$

then

$$\|f'\| \leqq \frac{8\sin(u/2)}{3+\cos u}. \tag{2.5}$$

The constant on the right side of (2.5) *is best.*

Observe that if we let $r=0$, hence $u=\pi$ by (2.3), then Theorem 2 reduces to Theorem 1 for complex-valued functions $f(t)$. Therefore Theorem 2 is a generalization of Theorem 1 for complex-valued functions.

In Section 3 we reformulate Theorem 2 in kinematic terms which will show where the peculiar right-hand sides of (2.4) and (2.5) come from.

3. Reformulation of Theorem 2 in terms of motions. Figure 1 shows the circle C_r of radius r, $0<r<1$, and the circle C_1 of radius 1, both with centers at the origin O. We identify letters such as $A, B, \cdots$, with the complex numbers that the corresponding points represent. Thus $B=1$, $E=r$, $D = r\cos(u/2)$, etc. Let $\Pi = ABA'$ denote the parabolic arc with vertex at B, axis OB, and such that Π is tangent to C_r at the points A, A'. Let $u = \measuredangle AOA'$ and let us determine the relation between the angle u and the radius r. Let the tangents to C_r at A and A' intersect in V. The point V is the inverse of D with respect to the circle C_r, and therefore $OV = (OE)^2/OD = r/\cos(u/2)$. From the parabola we see that the points V, B, D, and the point at infinity form a harmonic division, hence that B is the midpoint of DV. Therefore $B = 1 = \frac{1}{2}(r\cos(u/2) + r/\cos(u/2))$, whence we find that

$$r = \frac{4\cos(u/2)}{3+\cos u}, \qquad (0<u<\pi). \tag{3.1}$$

We now define a motion $f_1(\tau)$ of a point in the time interval $-\frac{1}{2} \leqq \tau \leqq \frac{1}{2}$, such that $f_1(\tau) = x_1(\tau) + iy_1(\tau)$ should describe the parabolic arc Π, with $f_1(-\frac{1}{2}) = A$, $f_1(\frac{1}{2}) = A'$, and such that $y_1(\tau)$ should be a *linear* function of τ. Evidently

$$y_1(\tau) = (2r\sin u/2)\tau, \qquad (-\tfrac{1}{2} \leqq \tau \leqq \tfrac{1}{2}).$$

To find $x_1(\tau)$ we observe that it must be of the form $x_1(\tau) = 1 - a\tau^2$. Since we want that $x_1(\pm\frac{1}{2}) = D = r\cos(u/2)$, we easily find that $a = 4(1-\cos u)/(3+\cos u)$. Therefore $x_1(\tau) = 1 - 4(1-\cos u)\tau^2/(3+\cos u)$. Changing variable by $\tau = t - (1/2)$, and setting $x_1(\tau) = x_0(t)$,

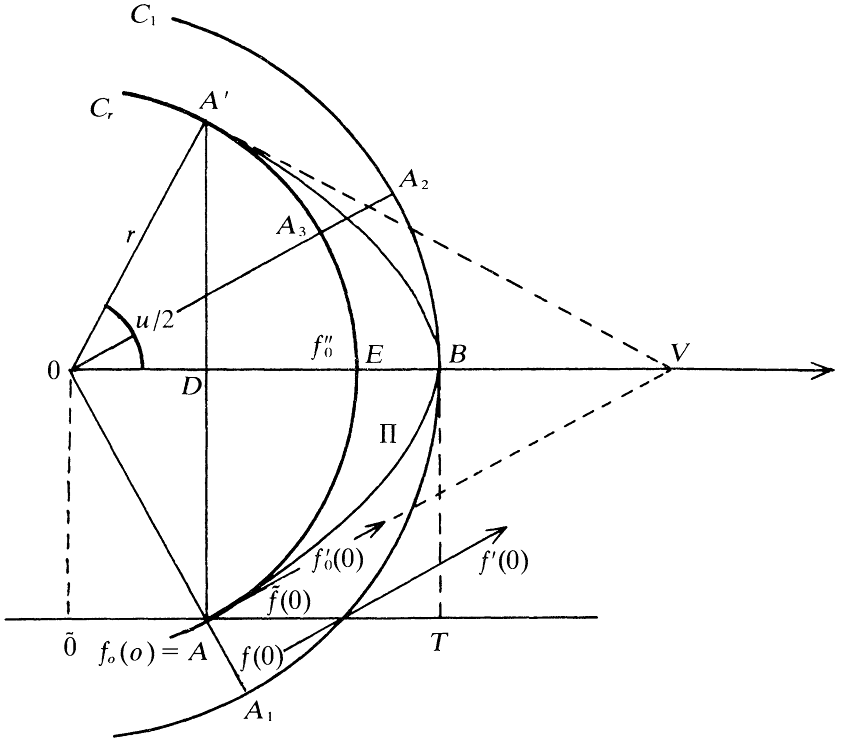


FIG. 1

$y_1(\tau) = y_0(t)$, $f_1(\tau) = f_0(t)$, we obtain the motion

$$f_0(t) = x_0(t) + iy_0(t), \qquad (0 \leqq t \leqq 1), \tag{3.2}$$

which describes the arc Π in the unit time-interval $0 \leqq t \leqq 1$. From previous results we find that

$$x_0(t) = 1 - \frac{1 - \cos u}{3 + \cos u}(2t-1)^2,$$
$$y_0(t) = \frac{2 \sin u}{3 + \cos u}(2t-1), \qquad (0 \leqq t \leqq 1). \tag{3.3}$$

We now extend the motion (3.2) *to all real values of t by means of the functional equation*

$$f_0(t+1) = e^{iu} f_0(t) \quad \textit{for all } t. \tag{3.4}$$

Observe the effect of this extension: Since $f_0(t)$ $(0 \leqq t \leqq 1)$ describes the arc Π, $f_0(t+1)$ $(0 \leqq t \leqq 1)$ will describe an arc Π_1 obtained by turning Π rigidly about O by an angle u. It follows that the path of the motion

$$f_0(t) = f_0(t;\, r) = x_0(t) + iy_0(t), \qquad (-\infty < t < \infty), \tag{3.5}$$

is an infinite succession of parabolic arcs $\ldots \Pi_{-1}, \Pi, \Pi_1, \Pi_2, \ldots$ obtained from Π by turning it about O by integral multiples of the angle u.

The relations (3.3) and (3.4) (or Fig. 1) imply that the motion $f_0(t)$ is continuous for all values of t and that it satisfies the inequalities

$$r \leqq |f_0(t)| \leqq 1 \quad \text{for all } t. \tag{3.6}$$

Again from (3.3), (3.4) (or Fig. 1) we easily find that also the velocity $f_0'(t)$ is continuous for all t, and that

$$\|f_0'\| = |f_0'(0)| = \frac{8\sin(u/2)}{3+\cos u}. \tag{3.7}$$

Similarly we find that the acceleration $f_0''(t)$ is piecewise constant and such that

$$f_0''(t) = -\frac{16\sin^2(u/2)}{3+\cos u}e^{i\nu u} \quad \text{if } \nu < t < \nu+1, \ (\nu \text{ integer}). \tag{3.8}$$

Therefore

$$\|f_0''\| = \frac{16\sin^2(u/2)}{3+\cos u}. \tag{3.9}$$

That the right sides of (3.7) and (3.9) are also the values of the constants in Theorem 2 is no accident, for they are the source of these constants. In terms of the motion (3.5) we may therefore restate Theorem 2 as follows.

THEOREM 2′. *If $f(t)$ is a motion satisfying the conditions*

$$r \leqq |f(t)| \leqq 1 \quad \textit{for all } t, \quad \|f''\| \leqq \|f_0''\|, \tag{3.10}$$

then

$$\|f'\| \leqq \|f_0'\|. \tag{3.11}$$

In this reformulation it is clear that $\|f_0'\|$ is the best constant in (3.11) because the motion $f_0(t)$ satisfies the conditions of Theorem 2′.

4. A proof of Theorem 1. Theorem 1 was established in [**5**] by means of a certain approximate differentiation formula. I have not succeeded to establish Theorem 2′ by a similar method. We record here the following direct proof of Theorem 1, because this approach requires only some geometric elaborations in order to furnish a proof of Theorem 2′.

Let the real-valued motion $f(t)$ satisfy the assumptions (1.1) of Theorem 1. We are to show that $-4 \leqq f'(t) \leqq 4$ for all t. Because of the translation-invariance of the conditions on f (if $f(t)$ satisfies them, then so does $f(t-c)$), it suffices to show that $-4 \leqq f'(0) \leqq 4$. Clearly $-1 \leqq f(0) \leqq 1$, but we may even assume that $0 \leqq f(0) \leqq 1$, or else we consider $-f(t)$. If $f'(0) = 0$, there is nothing to prove, so let $f'(0) \neq 0$. We may even assume that $f'(0) > 0$, for if $f'(0) < 0$, we consider the equally acceptable motion $f(-t)$. Let us therefore assume that

$$0 \leqq f(0) \leqq 1, \qquad f'(0) > 4 \tag{4.1}$$

and try to reach a contradiction. This we obtain as follows. In the interval $0 \leqq t \leqq 1/2$ we have, by (1.1), on the one hand that

$$\left| \int_0^t (t-v)f''(v)dv \right| \leqq 8\int_0^t (t-v)dv = 4t^2 \leqq 1. \tag{4.2}$$

On the other hand, by (4.1) and (4.2), and assuming that $0 < t \leqq 1/2$, we have

$$f(t) = f(0) + f'(0)t + \int_0^t (t-v)f''(v)dv > 0 + 4t - 1 \tag{4.3}$$

and in particular

$$f(\tfrac{1}{2}) > 0 + 4(1/2) - 1 = 1 \tag{4.3′}$$

showing that for $t = 1/2$ the point $f(t)$ has left the interval $[-1, 1]$, which is forbidden by (1.1).

REMARK. Although the above proof is complete, it is nevertheless useful for later applications to rephrase the inequalities (4.1) in terms of the extremizing motion appropriate for Theorem 1. If in the motion (3.2), (3.3), we set $r = 0$, hence $u = \pi$, we obtain the real-valued motion $f_0(t) = f_0(t; 0)$ defined in the time interval $0 \leqq t \leqq 1$ by the equations

$$\text{(4.4)} \qquad \begin{aligned} x_0(t) &= 4t - 4t^2, \\ y_0(t) &= 0, \qquad (0 \leqq t \leqq 1), \end{aligned}$$

which we extend to all real t, in view of (3.4), by the functional equation

$$\text{(4.5)} \qquad f_0(t+1) = -f_0(t).$$

Observing that $f_0(0) = 0$, $f_0'(0) = 4$, $f_0(\frac{1}{2}) = 1$, we may rewrite the inequalities (4.1) as

$$\text{(4.6)} \qquad f_0(0) \leqq f(0) \leqq f_0(\tfrac{1}{2}), \quad f_0'(0) < f'(0).$$

We see the following: the motion $f_0(t)$, which is maximally decelerated, just manages to turn around at the point $x = 1$ with $f_0'(\frac{1}{2}) = 0$. However, the motion $f(t)$, which starts ahead of $f_0(t)$, because of $f(0) \geqq f_0(0)$, with greater initial velocity, because of $f'(0) > f_0'(0)$, and no greater deceleration, because of $|f''(t)| \leqq 8$, must overrun the point $x = 1$, for $t = \frac{1}{2}$, as shown by (4.3′).

5. A proof of Theorem 2′ for the case when $\pi/2 \leqq u < \pi$. Fig. 1 illustrates this case. If OA intersects the unit circle in A_1, and if we turn the segment OAA_1 by an angle $\pi/2$ in the positive direction to obtain OA_3A_2, then no point of the closed curvilinear quadrilateral

$$\text{(5.1)} \qquad Q = AA_1A_2A_3$$

is to the left of the infinite vertical line AA'.

Let $f_0(t) = f_0(t; r)$ denote the special motion (3.2). Let $f(t)$ be any motion satisfying the conditions (3.10), and we are to show that it also satisfies (3.11). Evidently $\|f_0'\| = |f_0'(0)|$. We are to show that $|f'(t)| \leqq |f_0'(0)|$ for all t. It is clearly sufficient to show that $|f'(0)| \leqq |f_0'(0)|$. Let us assume that

$$\text{(5.2)} \qquad |f'(0)| > |f_0'(0)|$$

and try to reach a contradiction. We first turn the motion $f(t)$, i.e., replace $f(t)$ by $e^{i\alpha}f(t)$ (α real), so that the velocity vector $f'(0)$ becomes parallel to the vector $f_0'(0)$. We can also assume that (after the turning) the point $f(0)$ is below, or on, the infinite line OA_2. For, if this is not the case, we replace $f(t)$ by $-f(t)$. Finally, we may even assume that $f(0)$ is in the closed quadrilateral Q of (5.1). For, if this is not already the case, we can achieve it by reflecting the motion $f(t)$ in the line OA. Indeed, since the vector $f'(0)$ is already perpendicular to OA, it will remain so after a reflection in OA. The vectors $f'(0)$ and $f_0'(0)$ are now parallel. However, we can even assume that these vectors point in the same direction. For if this is not already the case, we can achieve this by replacing $f(t)$ by $f(-t)$, for this changes $f'(0)$ into $-f'(0)$.

After these preliminary changes we have the situation shown in Fig. 1: *The point $f(0)$ belongs to the quadrilateral Q and the vector $f'(0)$ is parallel to and has the same sense as* $f_0'(0)$. In order to reach a contradiction with (5.2), we draw through A the horizontal line AT and project orthogonally both motions $f(t)$ and $f_0(t)$ on this line, obtaining the motions $\tilde{f}(t)$ and $\tilde{f}_0(t)$, respectively. Observe the following:

1. $\tilde{f}_0(t)$ is in the interval $0 \leqq t \leqq 1$ a uniformly decelerated motion with deceleration $\|f_0''\|$. The motion starts from $\tilde{f}_0(0) = A$ with the initial velocity $|f_0'(0)|\cos\gamma$, where γ is the angle of $f_0'(0)$ with AT. Finally that $\tilde{f}_0(t)$ reaches T, for $t = \frac{1}{2}$, with velocity zero.

2. $\tilde{f}(t)$ leaves from the initial position $\tilde{f}(0)$, somewhere between A and T, with the positive initial velocity

$$\tilde{f}'(0) = |f'(0)|\cos\gamma > |f_0'(0)|\cos\gamma = \tilde{f}_0'(0)$$

in view of our assumption (5.2). Finally that

$$|\tilde{f}''(t)| \leqq |f''(t)| \leqq \|f_0''\|$$

by (3.10).

We see that we have before us the situation of the Remark of the last paragraph of Section 4, and may conclude that the point $\tilde{f}(\frac{1}{2})$ is on AT, but to the right of T. This contradicts the first assumption (3.10) which was left unchanged by our preliminary changes.

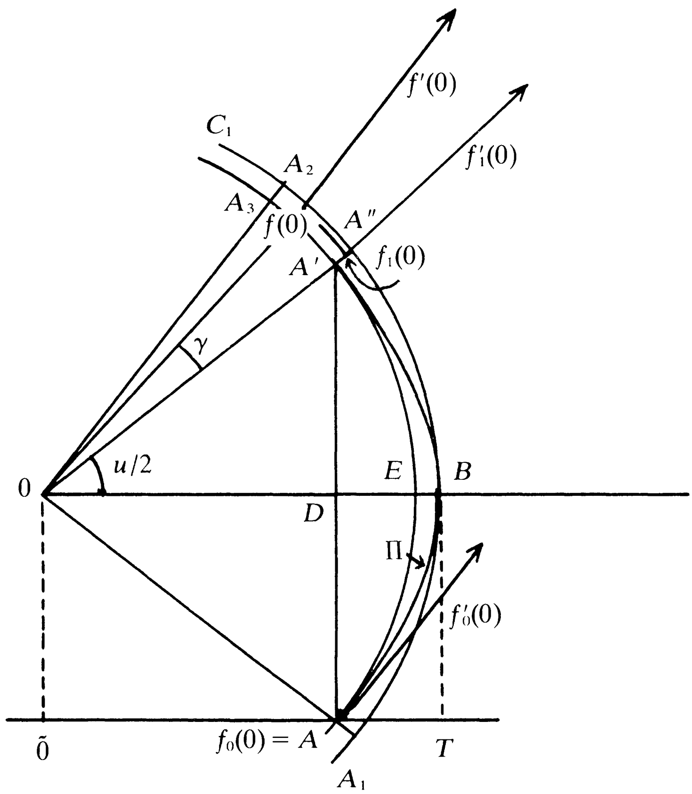


FIG. 2

6. A proof of Theorem 2′ for the case when $0 < u < \pi/2$. Fig. 2 shows the two circles C_r and C_1 as in Fig. 1. However, in the present case the $\frac{1}{4}$-ring $Q = AA_1A_2A_3$ contains the point A' on its boundary, somewhere between A and A_3. To establish Theorem 2′ by contradiction we assume again that (5.2) holds. We perform the similar turning and reflexions on $f(t)$ so as to have $f(0) \in Q$, while $f'(0)$ is parallel to and has the same orientation as $f_0'(0)$. Let OA' intersect C_1 in A'', and consider the curvilinear quadrilateral $Q' = AA_1A''A'$. If $f(0) \in Q'$ then projecting the motions $f(t)$ and $f_0(t)$ on the line AT will complete the proof by contradiction, as in the previous case. However, if the point $f(0)$ is to the left of the vertical line AA', then we cannot obtain the desired contradiction (see Fig. 2). Indeed, the point $\tilde{f}(0)$ now lies on AT to the *left* of A. This might give the motion $\tilde{f}(t)$ a chance of slowing down sufficiently so as to avoid overrunning the forbidden point T. This difficulty is overcome by one last rotation as follows: With center in O and radius equal to $|f(0)|$ we draw a clockwise circular arc $f(0)f_1(0)$ so that its endpoint $f_1(0)$ lies on the closed segment $A'A''$ (Fig. 2). Let

$f_1(0) = e^{-i\gamma}f(0)$. We replace the motion $f(t)$ by the new motion

$$f_1(t) = e^{-i\gamma}f(t).$$

We claim that the new motion $f_1(t)$ will provide the desired contradiction. Indeed, on projecting $f_1(t)$ on the line AT we obtain the motion $\tilde{f}_1(t)$ and it is clear that the initial position $\tilde{f}_1(0)$ is at A, or between A and T. Moreover, observe that the vector $f_1'(0)$ arises from $f'(0)$ by a clockwise turn through the angle γ, while

$$0 < \gamma \leqq \measuredangle A'OA_3.$$

It follows that

$$\tilde{f}_1'(0) > \tilde{f}'(0) > \tilde{f}_0'(0),$$

and our previous argument applies and shows that $\tilde{f}_1(1/2)$ has overrun the point T.

7. The nature of extremizing motions. We now assume that we have the equality sign (3.11) and that the extremum $\|f'\|$ is actually attained. This means that

$$|f'(t_0)| = \|f_0'\| \tag{7.1}$$

for some appropriate t_0. We may clearly assume that $t_0 = 0$. By appropropriate rotations and reflexions we may therefore assume that we have the following situation:

(α) The point $f(0)$ is in the quadrilateral $Q = AA_1A_2A_3$ of Fig. 1 in the case of §5, or else in $Q' = AA_1A''A'$ of Fig. 2 in the case of §6.

(β) The initial velocity vectors of the motions f and f_0 agree in size and direction so that

$$f'(0) = f_0'(0). \tag{7.2}$$

For convenience we think of AT as a real axis, with origin at $\tilde{O}$, so that $\tilde{f}(t)$ and $\tilde{f}_0(t)$ are now real-valued functions. Whether u is $\geqq \pi/2$, or else $< \pi/2$, we have that

$$\tilde{f}_0(0) \leqq \tilde{f}(0) \tag{7.3}$$

and

$$\tilde{f}_0'(0) \leqq \tilde{f}'(0). \tag{7.4}$$

Observe that we may well have the inequality sign in (7.4) due to the last rotation by γ in Fig. 2. We also know that $\tilde{f}_0''(t)$ is constant and $= -\|f_0''\|$ in the interval $0 \leqq t \leqq 1$. By Taylor's formula (4.3) for $t = 1/2$ we obtain

$$\tilde{f}_0(1/2) = \tilde{f}_0(0) + \tilde{f}_0'(0)\tfrac{1}{2} - \int_0^{1/2} (\tfrac{1}{2} - v)\|f_0''\|\, dv = 1. \tag{7.5}$$

From $\|f\| \leqq 1$ we also conclude that

$$\tilde{f}(1/2) = \tilde{f}(0) + \tilde{f}'(0)\tfrac{1}{2} + \int_0^{1/2} (\tfrac{1}{2} - v)\tilde{f}''(v)dv \leqq 1. \tag{7.6}$$

If we subtract (7.5) from (7.6) we see that

$$(\tilde{f}(0) - \tilde{f}_0(0)) + \tfrac{1}{2}(\tilde{f}'(0) - \tilde{f}_0'(0)) + \int_0^{1/2} (\tfrac{1}{2} - v)(\tilde{f}''(v) + \|f_0''\|)dv \leqq 0.$$

However, the first two terms as well as the integrand are non-negative by (7.3), (7.4) and we conclude that

$$\tilde{f}(0) = \tilde{f}_0(0), \quad \tilde{f}'(0) = \tilde{f}_0'(0), \tag{7.7}$$

as well as

$$\tilde{f}''(v) = -\|f_0''\| \quad \text{if } 0 \leqq v \leqq 1/2. \tag{7.8}$$

From these relations we draw several consequences:

1. From the first relation (7.7) we conclude that $f(0) = A$ in the case of Fig. 1, or else $f(0) = A$ or $= A'$ in Fig. 2.

2. From (7.8) and $\|f''\| \leqq \|f_0''\|$ we infer that $\operatorname{Im} f''(v) = 0$ if $0 \leqq v \leqq 1/2$, and therefore that

$$f''(v) = -\|f_0''\| \quad \text{if } 0 \leqq v \leqq 1/2. \tag{7.9}$$

By (7.7), (7.8), we must have the equality sign in (7.6) as seen if we compare (7.5) with (7.6). Thus $\tilde{f}(\frac{1}{2}) = 1$ and therefore

$$f(1/2) = B. \tag{7.10}$$

This last equation implies that

$$\operatorname{Re} f'(\tfrac{1}{2}) = 0. \tag{7.11}$$

The three conditions (7.9), (7.10) and (7.11) amply imply that

$$f(t) = f_0(t) \text{ in the interval } 0 \leqq t \leqq 1/2, \tag{7.12}$$

or perhaps

$$f(t) = \overline{f_0(t)} = x_0(t) - iy_0(t) \text{ in } (0, 1/2).$$

Replacing $f(t)$ by $f(-t)$ we establish that

$$f(t) = f_0(t) \text{ also in the interval } -1/2 \leqq t \leqq 0, \tag{7.13}$$

or perhaps $f(t) = \overline{f_0(t)}$ in $(-1/2, 0)$. This completes a proof of the following:

THEOREM 3. *Let $f(t)$ satisfy the assumptions* (3.10) *of Theorem* 2'. *If for some $t = t_0$ we have*

$$|f'(t_0)| = \|f_0'\|, \tag{7.14}$$

then

$$f(t) = e^{i\alpha} f_0(\varepsilon(t - t_0)) \text{ if } t_0 - \tfrac{1}{2} \leqq t \leqq t_0 + \tfrac{1}{2}, \tag{7.15}$$

for an appropriate real α and $\varepsilon = \pm 1$.

Outside of the interval $(t_0 - \frac{1}{2}, t_0 + \frac{1}{2})$ there are many possibilities of extending (7.15), one of which is to use (7.15) for all real t. To see this, let us extend the motion

$$f_0(t) \, (-\tfrac{1}{2} \leqq t \leqq \tfrac{1}{2})$$

in a way different from the obvious one. By (3.3) we obtain

$$f_0'(1/2) = iy_0'(1/2) = i4 \sin u/(3 + \cos u) = i\omega.$$

Now $f_0(1/2) = 1$, $f_0(-1/2) = e^{-iu}$, and

$$f(t) = \begin{cases} f_0(t) & \text{if } \ -1/2 \leqq t \leqq 1/2, \\ e^{i\omega(t-1/2)} & \text{if } \ t > 1/2, \\ e^{i\omega(t+1/2)} e^{-iu} & \text{if } \ t < -1/2, \end{cases}$$

is found to be an acceptable extension. The reason is that $\omega^2 < \|f_0''\|$ is a consequence of $0 < u < \pi$, as readily verified.

In contrast to Theorem 3 there very likely exist motions $f(t)$, satisfying the conditions (3.10) of

Theorem 2′, such that $\|f'\| = \|f_0'\|$, while

(7.16) $$|f'(t)| < \|f_0'\| \quad \text{for all real } t.$$

For a discussion and construction of such *extremizing functions in the weak sense* for Šilov's Theorem 2 of [**5**, p. 131] see [**5**, §9]. A construction of extremizing motions for our Theorem 2′ satisfying (7.16) would seem more difficult due to our requirement that $|f(t)| \geqq r$ for all real t.

8. Landau's problem in bounded continua. We begin with a few definitions.

Definition 1. We consider motions $f(t)$ $(t \in R)$ in the complex plane $\mathbb{C}$ satisfying the following conditions:

(8.1) $$f(t) \textit{ is bounded and continuously differentiable on } R,$$

(8.2) $$f'(t) \textit{ satisfies a Lipschitz condition with the least constant } A, \ (A > 0),$$

i.e., the inequality $|f'(t_1) - f'(t_2)| \leqq A\,|t_1 - t_2|$ holds for all real t_1, t_2, and no smaller constant A', in place of A, will do. It follows that $f''(t)$ exists almost everywhere (a.e.) and that its norm $\|f''\|$ = essential supremum of $|f''(t)|$ on R, has the value

(8.3) $$\|f''\| = A.$$

We denote the class of such motions by the symbol $\mathscr{A}$ and call its elements *admissible.*

Definition 2. Let K be a bounded continuum in $\mathbb{C}$, i.e., a bounded, closed and connected set. We assume that *there exists a motion.*

(8.4) $$f(t) \in \mathscr{A} \textit{ such that } f(t) \in K \textit{ for all real } t.$$

We then say that K is an *admissible continuum.* We denote the class of motions $f(t)$ satisfying (8.4) by the symbol $\mathscr{A}(K)$ and say that $f(t)$ is *admissible in K.*

An example of an admissible continuum is the ring $K_r = \{z\,;\, r \leqq |z| \leqq 1\}$ of Theorem 2′ (of Section 3), because $f_0(t)$, defined by (3.2)–(3.4), clearly belongs to $\mathscr{A}(K_r)$. Also the circular motion $f(t) = \mathrm{Re}^{i\omega t}$ $(R > 0,\ \omega$ real $\neq 0)$ is admissible and shows that its path $|z| = R$ is an admissible continuum.

Definition 3. *For every admissible $f(t)$ we define the functional $\mathscr{F}(f)$*

(8.5) $$\mathscr{F}(f) = \|f'\|/\sqrt{\|f''\|}.$$

Two obvious but relevant properties of this functional are the following:

(i) *If c and d are real, $c \neq 0$. Then*

(8.6) $$\mathscr{F}(f(ct + d)) = \mathscr{F}(f(t))$$

which means that $\mathscr{F}(f(t))$ is invariant if we change scale and origin in t. Indeed, observe that $\|D_t f(ct + d)\| = |c|\,\|f'\|$ and

(8.7) $$\|D_t^2 f(ct + d)\| = c^2 \|f''\|,$$

and now (8.5) implies (8.6).

(ii) Similarly we find that

(8.8) $$F(cf(t)) = \sqrt{c}\,\mathscr{F}(f(t)) \qquad (c > 0).$$

Definition 4. *If K is an admissible continuum, we define the Landau constant $\mathscr{L}(K)$ by*

(8.9) $$\mathscr{L}(K) = \sup \mathscr{F}(f) \quad \textit{for } f \in \mathscr{A}(K).$$

Conceivably we may have $\mathscr{L}(K)=\infty$, but we shall see below (Corollary 2) that it is always finite.

From the definition of $\mathscr{L}(K)$ as a supremum we immediately obtain the following *monotonicity property*: *If K' and K'' are admissible continua, then*

$$K' \subset K'' \text{ implies that } \mathscr{L}(K') \leqq \mathscr{L}(K''). \tag{8.10}$$

Our present generalization of Landau's problem is as follows: *To determine the value of $\mathscr{L}(K)$ for a prescribed admissible continuum K.*

We begin our discussion of this problem by stating the equivalence of two different ways of solving it. Let K be an admissible continuum and suppose that we have a special motion

$$f_0(t) \in \mathscr{A}(K) \tag{8.11}$$

enjoying the

Property A. *If*

$$f(t) \in \mathscr{A}(K) \text{ and } \|f''\| \leqq \|f_0''\| \tag{8.12}$$

then

$$\|f'\| \leqq \|f_0'\|. \tag{8.13}$$

This we wish to compare with

Property B. *The same motion $f_0(t)$, satisfying* (8.11) *is such that*

$$\mathscr{L}(K) = \mathscr{F}(f_0). \tag{8.14}$$

Theorem 4. *The two Properties* A *and* B *of $f_0(t)$ are equivalent.*

We omit the easy proofs of both implications.

Let us apply Theorem 4 to the ring

$$K_r = \{z\,;\, r \leqq |z| \leqq 1\} \quad (0 \leqq r < 1), \tag{8.15}$$

which is the continuum appearing in Theorem 2′. Theorem 2′ shows that K_r and the motion $f_0(t)$ defined by (3.2), (3.3) and (3.4), do have the Property A. Moreover, from (3.7) and (3.9) we find that

$$\mathscr{F}(f_0) = \frac{2}{\sqrt{3+\cos u}} \qquad (0 < u \leqq \pi). \tag{8.16}$$

By Theorem 4 we conclude that our Theorem 2′ is fully equivalent to the

Corollary 1. *For the ring* (8.15) *we have*

$$\mathscr{L}(K_r) = \mathscr{F}(f_0) = \frac{2}{\sqrt{3+\cos u}}, \qquad (0 < u \leqq \pi).^{\dagger} \tag{8.17}$$

† We refer to Figure 1 and let $B' = -1$, so that $B'B$ is a diameter of the circle C_1. Let Γ denote the ellipse having the segment $[B', E]$ as major axis and having a focus at the origin O. (The reader is asked to draw a sketch of the ellipse Γ and to observe that $\Gamma \subset K_r$.) By Kepler's first law there is a motion $f_K(t)$ of a material particle P moving along Γ under the action of a force directed towards O and proportional to $(OP)^{-2}$. We refer to $f_K(t)$ as a *Keplerian motion* along Γ. By Kepler's second law (the law of areas) we find that

$$\|f_K'\| = |f_K'(0)|, \quad \text{while evidently } \|f_K''\| = |f_K''(0)|,$$

where we assume that the *perihelion* E is reached at the time $t=0$. From the information contained in the first four sections of Chapter One of Harry Pollard's *Mathematical introduction to celestial mechanics*, Prentice-Hall, 1966,

Here r and u are connected by the relation (2.3), or

(8.18) $$r = \frac{4\cos(u/2)}{3+\cos u}, \qquad (0 < u \leqq \pi, \text{ hence } 0 \leqq r < 1).$$

Let

(8.19) $$K(R) = \{z\colon |z| \leqq R\}$$

denote the circular disk of radius R. In our notation (8.15) we may also write $K(1) = K_0$. Now (8.17) also holds for $r = 0$, hence $u = \pi$. We thus obtain that

(8.20) $$\mathscr{L}(K(1)) = \sqrt{2}.$$

Clearly $f(t) \in \mathscr{A}(K(R))$ if and only if $f(t)/R \in \mathscr{A}(K(1))$. By (8.8) we obtain $\mathscr{F}(f(t)/R) = R^{-1/2}\mathscr{F}(f)$, and equating the suprema of both sides we obtain, by (8.9) and (8.20), that $\sqrt{2} = R^{-1/2}\mathscr{L}(K(R))$, whence

(8.21) $$\mathscr{L}(K(R)) = \sqrt{2R}.$$

In words: *The Landau constant of a circular disk is equal to the square root of its diameter.* Theorem 5 below will generalize this to arbitrary convex and bounded K.

An immediate consequence of (8.21) is

COROLLARY 2. *For every admissible continuum K the Landau constant $\mathscr{L}(K)$* (Definition 4) *is well defined and finite.*

For if K is admissible and R is sufficiently large then $K \subset K(R)$. The monotonicity property (8.10), and (8.21), show that $\mathscr{L}(K) \leqq \sqrt{2R}$, and therefore $\mathscr{L}(K)$ is finite.

Our last remark of the present section is as follows. In Corollary 1 we have assumed that $r < 1$, hence $u > 0$. Let us show that it holds also for $r = 1$, hence $u = 0$. We state this as

COROLLARY 3. *For the unit circumference $K_1 = \{z;\ |z| = 1\}$ we have*

(8.22) $$\mathscr{L}(K_1) = \mathscr{F}(e^{it}) = 1.$$

Proof. If $0 < r < 1$ then $K_1 \subset K_r$, and the property (8.10) shows that

$$\mathscr{L}(K_1) \leqq \mathscr{L}(K_r) = 2/\sqrt{3+\cos u} \to 1 \text{ as } u \to 0+.$$

Therefore $\mathscr{L}(K_1) \leqq 1$. However, the motion $f(t) = e^{it}$ belongs to $\mathscr{A}(K_1)$, while we easily find that $\mathscr{F}(e^{it}) = 1$. This gives the reverse inequality $\mathscr{L}(K_1) \geqq 1$ and establishes Corollary 3.

For a circular disk of radius R we found the relation (8.21). We conclude our discussion with the following general theorem which we state without proof:

---

we readily find that

(*) $$\mathscr{F}(f_K) = |f_K'(0)| \cdot |f_K''(0)|^{-\frac{1}{2}} = \sqrt{2r/(1+r)}.$$

By Corollary 1, if $0 < r < 1$, we must have the inequality

(**) $$\sqrt{2r/(1+r)} < 2/\sqrt{3+\cos u}.$$

This is readily verified directly: Writing $x = \cos(u/2)$, and using the relation (8.18), we easily find that (**) amounts to the obvious inequality $(1-x)(1+x+2x^2) > 0$ if $0 < x < 1$. Observe that if $r = 1$, hence $u = 0$, then both sides of (**) become equal, which is allright, because in this case $f_K(t) = e^{it}$ is an acceptable Keplerian motion. The value of $\mathscr{F}(f_K)$ as given by (*) was kindly checked by A. A. Deprit, Professor of Celestial Mechanics at the University of Cincinnati, visiting our MRC this year.

THEOREM 5. *If K is compact and convex, then $\mathscr{L}(K) = \sqrt{D}$, where D is the set-theoretic diameter of K.*

Sponsored by the United States Army under Contract No. DAAG29–75–C–0024.

References

1. A. S. Cavaretta, Jr., An elementary proof of Kolmogorov's theorem, this MONTHLY, 21 (1974) 480–486.

2. C. de Boor and I. J. Schoenberg, Cardinal interpolation and spline functions VIII. The Budan-Fourier theorem for splines and applications, MRC Technical Summary Report # 1546. Appeared in "Spline Functions," Proc. of an Intern. Symp. held at Karlsruhe, May 20–23, 1975, Lecture Notes in Math. No. 501, Springer–Verlag, pp. 1–77.

3. A. Kolmogorov, On inequalities between the upper bounds of the successive derivatives of an arbitrary function on an infinite interval, Amer. Math. Soc. Transl., Series 1, 2 (1962) 233–243. Appeared in Russian in 1939.

4. E. Landau, Einige Ungleichungen für zweimal differenzierbare Funktionen, Proc. London Math. Soc., (2) 13 (1913) 43–49.

5. I. J. Schoenberg, The elementary cases of Landau's problem of inequalities between derivatives, this MONTHLY, 80 (1973) 121–158.

6. ———, Cardinal spline interpolation, CBMS vol. 12, SIAM, Philadelphia, 1973.

MATHEMATICS RESEARCH CENTER, UNIVERSITY OF WISCONSIN, MADISON, WI 53706.

WARING'S PROBLEM MOD n

CHARLES SMALL

0. Introduction. Our aim in this paper is to compute $g(k, n)$, which is by definition the smallest r such that every integer is a sum of r kth powers mod n. (*To avoid trivialities, we assume $k > 1$ and $n > 1$ throughout.*) En route, we survey some aspects of the classical Waring problem (i.e., for natural numbers rather than numbers mod n), in Section 1 and Section 7. We get complete results for $g(k, n)$ for $k = 2$ in Section 3, and for $k = 3$ in Section 5, and partial results for larger k in Section 6. The hard kernel of the problem, which we do not succeed in cracking, is the precise computation of $g(k, p)$ for p prime, $k \mid p-1$, $3 < k \leqq (p-1)/3$, $p < (k-1)^4$.

Our method to study $g(k, n)$ is as follows. First we look (in Section 4) mod primes p; here a theorem of Vosper gives a good upper bound. Then standard lifting techniques (described in Section 2) are applied, to take us from primes p to prime powers p^e. Finally, the Chinese Remainder Theorem allows us to pass from a prime-power modulus to an arbitrary one.

To study sums of kth powers mod n is, of course, to study congruences of the form

$$(*) \qquad \sum_{i=1}^{r} x_i^k \equiv b \bmod n.$$

Ideally, one would like a formula for the number of solutions to (*), as a function of n, k, r and b. The question studied here is weaker: how big must r be, as a function of n and k, to ensure the existence of at least one solution for each b? The general problem was studied by Hardy and Littlewood, using analytic methods, in connection with Waring's problem, **[8a]**. It is also a starting point for the theory of rational points of varieties defined over finite fields, which "began" with **[22]**, and reached a high point with the recent completion, by Deligne, of the proofs of the Weil conjectures. In contrast with all of this, our methods of attack on the weaker question are completely elementary.

The problem mod n is much easier, too, than the "global" problem, and accordingly our arguments involve none of the difficult machinery with which Waring's problem (over $\boldsymbol{Z}$) is

customarily attacked. In Section 7 we investigate some possible applications of our results mod n to outstanding questions in the global problem; unfortunately, we are able to show that all are fruitless, although in one case this is itself a non-trivial result (cf. 7.9).

Added in proof, October 1976: In fact, the techniques of this paper lead also to an explicit solution for fifth powers, analogous to the theorems for squares and cubes given in Sections 3 and 5, and to a nearly-complete solution for fourth powers. The solution for $k = 5$ leads in turn to a complete solution in principle for all odd k, in the following sense: we prove a theorem which first produces, for each fixed odd k, an *explicit finite* list n_i, $i \in I$, and then computes $g(k, n)$ for *all* n explicitly in terms of the $g(k, n_i)$, $i \in I$. (The cardinality of I is $s + t + 2\Sigma e_i$, where $k = p_1^{e_1} \cdots p_s^{e_s}$ is the prime factorization of k and t is the number of primes q such that $q \equiv 1 \pmod{2p_i}$ for some i and $q < (k-1)^4$; these primes q are among the n_i.) The problem is subtler when k is even, and the difficulty is already apparent in the case $k = 4$. The first obstacle, for even k, is to determine the primes p such that -1 is a sum of fewer than $g(k, p)$ kth powers modulo p. Details of all these additional results will appear shortly, as an Addendum to the present paper, in a subsequent issue of this MONTHLY.

1. Waring's problem. In this section we survey briefly the history of Waring's problem. (For another version, both deeper and broader, see [**6**], and for something in between see [**17a**].) This is not really necessary for the subsequent material on the situation mod n. We give it for two reasons: to provide general background and set our problem in context, and to make publicity for the "recent" nearly-complete results on Waring's original problem, which seem to be widely unknown, although they go back to the mid 1930's.

In 1770, Edward Waring published the assertion that every positive integer is the sum of 4 squares, 9 cubes, 19 biquadrates (fourth powers), "and so on." The assertion for squares is much older; according to Ellison [**6**] it goes back to Fermat (1640), according to Hardy and Wright [**9**] it goes back to Bachet (1621), and some authors trace it to Diophantus (3rd century A.D.). In any event, it was first proved by Lagrange, in 1770. (For relevant quotations from Fermat, Waring, Lagrange, etc. see [**1**].) The general theorem was not proved until 1909, when Hilbert showed that given a positive integer k, every positive integer is a sum of r kth powers of positive integers, for some *r depending only on k*. The smallest such r is traditionally called $g(k)$. Proofs of Hilbert's theorem can be found in [**14**], [**6**] and [**3a**].

Elementary texts on number theory, when they discuss Waring's problem, usually end the story at this point, perhaps with a sentence to the effect that Hilbert's proof is, unfortunately, purely existential, and does not determine the values of $g(k)$. It seems to be widely unknown that in fact $g(k)$ can be computed explicitly, for all $k \neq 4$. The theorem which computes $g(k)$ is due in part to many people, though perhaps chiefly to Dickson, circa 1936, based on work of Vinogradov and Hardy and Littlewood. (The case most recently settled is $k = 5$; this is due to Chen [**2**], 1964.)

In a sense, the key to the theorem in question is an observation first made by Euler. Given k, one divides 3^k by 2^k, to obtain integers q and r satisfying $3^k = q \cdot 2^k + r$, $0 \leqq r < 2^k$; in other words, $q = [(3/2)^k]$. Now let $l = q \cdot 2^k - 1$. Since $l < 3^k$, the only kth powers which can occur when we write l as a sum of kth powers are 1^k and 2^k; and since $l = (q-1) \cdot 2^k + (2^k - 1) \cdot 1^k$, it is clear that l is the sum of $q - 1 + 2^k - 1 = q + 2^k - 2$ kth powers and of no fewer. This shows that if we define $\bar{g}(k) = q + 2^k - 2$, where $q = [(3/2)^k]$ as above, then $g(k) \geqq \bar{g}(k)$ for all k. *It is not known whether this inequality is ever strict.* What *is* known can be summarized as follows:

THEOREM. *Let $k \neq 4$ be a positive integer, and put $q = [(3/2)^k]$, $q' = [(4/3)^k]$. Then*:

CASE 1. *If $r + q \leqq 2^k$ then $g(k) = \bar{g}(k)$.*

CASE 2. *If $r + q > 2^k$, then*

$$g(k) = \begin{cases} \bar{g}(k) + q' - 1 & \text{if} \quad 2^k < qq' + q + q' \\ \bar{g}(k) + q' & \text{if} \quad 2^k = qq' + q + q'. \end{cases}$$

(One sees easily that $2^k \leqq qq' + q + q'$ in any case.) For $k = 2, 3, 5, 6, 7, 8, 9, \ldots$ the theorem gives $g(k) = 4, 9, 37, 73, 143, 279, 548, \ldots$.

It is not known whether Case 2 really occurs. In **[18]** it is shown that all $k \leqq 200,000$ fall into Case 1, and in **[12a]** it is shown that at most finitely many k fall into Case 2.

For $k = 4$ we have $g(4) \geqq \bar{g}(4) = 19$ as above, and $g(4) \leqq 22$, **[19]**. There seems to be little doubt that additional computation will eventually prove $g(4) = 19$. Proof of the conjecture that $g(k) = \bar{g}(k)$ for all k, i.e., that Case 2 never occurs — the so-called "ideal Waring theorem" — seems more remote.

Except for the remark that $g(k) \geqq \bar{g}(k)$, proofs of the theorems described above are difficult. The determination of $g(k)$, for example, depends upon analytic methods of Vinogradov which give upper bounds for $G(k)$, defined as the smallest r such that every positive integer, *with perhaps finitely many exceptions*, is a sum of r kth powers. (We shall have more to say about $G(k)$ in Section 7.) A vast literature has grown up around these and related questions, under the general heading of "additive number theory." For a systematic account of the general theory, with extensive references and bibliography through 1956, see **[15]**. (**[6]** has a good list of more recent references on Waring's problem *per se*.)

Waring-type questions arise in arbitrary commutative rings; see **[10]**. There are even a few results in non-commutative situations; e.g. **[7]**.

2. A lifting lemma. The following result will help us get information mod prime powers p^e from information mod p. It is part of a standard result. We include the proof because it is easy and pretty, though we shall quote a more general form below (2.3).

2.1 *Let p be a prime and k a positive integer not divisible by p. Assume that in $\boldsymbol{Z}/(p)$, every element is a sum of r k-th powers, and 0 is nontrivially a sum of r k-th powers. Then, for all $e \geqq 1$, every element of $\boldsymbol{Z}/(p^e)$ is a sum of r k-th powers.*

Proof. Let $f(X_1, \ldots, X_r) = X_1^k + \cdots + X_r^k$. Given $a \in \boldsymbol{Z}$, we have a solution $(x_1, \ldots, x_r) \in \boldsymbol{Z}^r$ to the congruence $f(X_1, \ldots, X_r) \equiv a \mod p$, and $p \nmid x_i$ for at least one index i. Assume without loss of generality that $p \nmid x_1$ and put $g(X) = f(X, x_2, \ldots, x_r) - a$. Then $g(x_1) \equiv 0 \mod p$, and we proceed by induction : assume $g(x_1) \equiv 0 \mod p^n$ for some $n \geqq 1$, say

$$(*) \qquad g(x_1) = p^n b, \qquad b \in \boldsymbol{Z}.$$

We claim $g(x_1 + cp^n) \equiv 0 \mod p^{n+1}$, for some suitably chosen c. (Clearly this completes the proof.) For we have

$$g(x_1 + cp^n) = g(x_1) + \sum_{i=1}^{k} \binom{k}{i} x_1^{k-i} c^i p^{ni}$$

(analysts will see this as a Taylor expansion, and algebraists will recognize the binomial theorem), so that

$$(**) \qquad g(x_1 + cp^n) \equiv g(x_1) + ckx_1^{k-1} p^n \mod p^{n+1}.$$

But $p \nmid kx_1^{k-1}$ since $p \nmid k$ and $p \nmid x_1$. Thus kx_1^{k-1} is a unit mod p, and we can choose c to get $ckx_1^{k-1} \equiv -b \mod p$, so that $p^n(ckx_1^{k-1} + b) \equiv 0 \mod p^{n+1}$. But $p^n(ckx_1^{k-1} + b)$ is just $g(x_1 + cp^n)$ mod p^{n+1}, as one sees by combining (*) and (**). Done!

Now let $t(k, n)$ denote the smallest s such that 0 is nontrivially a sum of s kth powers mod n. (*Exercise*: $t(k, n) \neq 1 \Leftrightarrow n$ is square-free.) Note that $t(k, n) \leqq 2$ if k is odd $(1^k + (-1)^k = 0)$, and $t(k, n) \leqq g(k, n) + 1$ in any case (for a similar reason). The computation of $t(2, p)$ is well known, and is recalled below (3.3).

2.2 COROLLARY. *If p is prime and $p \nmid k$, then*

$$g(k,p) \leqq g(k,p^e) \leqq \max(g(k,p), t(k,p)) \leqq g(k,p)+1$$

for all $e \geqq 1$.

Proof: The first inequality is clear: anything true mod p^e is even truer mod p. The second follows from 2.1, and the third follows from $t(k,p) \leqq g(k,p)+1$.

REMARK: A very similar proof of 2.2 can be given, by taking advantage of the fact that $Z/(p^e)$ is a complete local ring with residue class field $Z/(p)$, and applying Hensel's lemma.

To handle the case when $p \mid k$, for example to compute $g(p,p^e)$, we need a stronger version of 2.1:

2.3 *Let p be a prime, let $f(X_1,\dots,X_r) = X_1^k + \cdots + X_r^k - a$ $(a \in Z)$, and suppose that $f(X_1,\dots,X_r) \equiv 0 \bmod p^n$ has a primitive solution $x_1,\dots,x_r$ $(x_i \in Z)$. ("Primitive" means $p \nmid x_j$ for at least one index j. It is automatic if $p \nmid a$.) Assume further that $n > 2c$, where p^c $(c \geqq 0)$ is the highest power of p which divides k (i.e., $p^c \| k$). Then $f(X_1,\dots,X_r) \equiv 0 \bmod p^e$ has solutions for all $e \geqq 1$.*

For a proof (of a more general theorem than 2.3) see **[17]**, Theorem 1, p. 14. (The statement of that theorem should have $1 \leqq j \leqq m$ and $0 \leqq 2k < n$.) Note that 2.1 is a special case, with $c = 0$ and $n = 1$, of 2.3. 2.3 allows us to improve congruences even in the unlucky case when $p \mid k$, but we have to start higher.

2.4 COROLLARY. *If p is prime and $p^c \| k$ $(c > 0)$, then*

$$g(k,p) \leqq g(k,p^2) \leqq \cdots \leqq g(k,p^{2c+1}) = g(k,p^{2c+i})$$

for all $i \geqq 1$.

Proof: It suffices to note that $t(k,p^e) = 1$ when $e > 1$ (compare the Exercise above): since $k > 1$, we can choose $m < e$ so that $mk \geqq e$; then $p^m \not\equiv 0$ but $(p^m)^k \equiv 0$, mod p^e.

When $k = p$, 2.4 gives $g(p,p) \leqq g(p,p^2) \leqq g(p,p^3) = g(p,p^{2+i})$ for all $i \geqq 1$. Clearly $g(p,p) = 1$: everything satisfies $x^p \equiv x \bmod p$. It is almost as clear that $g(p,p^2) > 1$; the reader can either supply a proof, or look ahead to 5.2. Finally, in Section 7 we will show that $g(p,p^2) = g(p,p^3)$ for odd primes p. Putting these remarks together, we see that when $k = p$ is an odd prime, 2.4 can be strengthened to: $1 = g(p,p) < g(p,p^2) = g(p,p^{2+i})$ for all $i \geqq 0$. (Note however that if $p \mid k$ but $p \neq k$ we need not have $g(k,p^2) = g(k,p^3)$: the sixth powers mod 27 are $\{0,1,10,19\}$, so that $g(6,9) = 8 < 9 = g(6,27)$.)

3. Sums of squares mod n. In this section we prove the following theorem, essentially due to Gauss:

3.1 *Let $r \geqq 1$ and $n > 1$ be integers. Then every element of $Z/(n)$ is a sum of r squares if and only if one of the following holds*:

(1) $r = 1$ *and* $n = 2$;

(2) $r = 2$ *and* $4 \nmid n$ *and* $p^2 \nmid n$ *for all primes* $p \equiv 3 \bmod 4$;

(3) $r = 3$ *and* $8 \nmid n$;

(4) $r \geqq 4$ (*and no condition on* n).

First, some lemmas:

3.2 *Let F be a finite field. Then every element of F is a sum of two squares. If* char $F = 2$ *then every element is actually a square; if* char $F \neq 2$ *then exactly half of the non-zero elements are squares.*

Proof: In characteristic 2, $x \mapsto x^2$ is injective. Hence, since F is finite, it is also surjective. So

assume char $F \neq 2$. The group $F^{\cdot}$ of non-zero elements is cyclic and of even order; say $F^{\cdot} = \{g, g^2, \ldots, g^{2m} = 1\}$. Clearly, exactly half of these are squares, viz. g^{2i}, $i = 1, \ldots, m$. Suppose $z \in F^{\cdot}$ is not a square, and let $S = \{x^2 \mid x \in F^{\cdot}\}$ and $T = \{z - y^2 \mid y \in F^{\cdot}\}$. S and T have the same number of elements, namely half the number of elements of $F^{\cdot}$. Since $z \notin S$ and $z \notin T$, it follows that $S \cap T$ is non-empty. Hence z is a sum of two squares.

REMARK: This proof, modified slightly, shows more generally that over a finite field F of characteristic $\neq 2$, every binary quadratic form is universal: given non-zero elements $a, b, c \in F$, there exist $x, y, \in F$ with $ax^2 + by^2 = c$. (Put $S = \{ax^2 \mid x \in F\}$, $T = \{c - by^2 \mid y \in F\}$; then S and T each have $m+1$ elements, where F has $2m+1$ elements, so that $S \cap T \neq \varnothing$.) In another direction, 3.2 generalizes from squares to kth powers: see 4.1 below.

3.3 *Let p be an odd prime. Then*:

(i) -1 *is a square* $\bmod p \Leftrightarrow p \equiv 1 \bmod 4$;

(ii) *if* $p \equiv 3 \bmod 4$, *then* 0 *is not nontrivially a sum of two squares* $\bmod p$, *and consequently p is not a sum of two squares* $\bmod p^2$;

(iii)
$$t(2,p) = \begin{cases} 2 & \text{if} \quad p \equiv 1 \bmod 4 \\ 3 & \text{if} \quad p \equiv 3 \bmod 4. \end{cases}$$

Proof: (i) If $p \equiv 1 \bmod 4$ write $p - 1 = 4n$. The group of units of $Z/(p)$ is cyclic of order $4n$, so that if g is a generator we have $(g^n)^2 = -1$. Conversely if $x^2 = -1$ then x has order 4, so $4 \mid p-1$.

(ii) If x, y are integers with $p \equiv x^2 + y^2 \bmod p^2$ we cannot have both $p \mid x$ and $p \mid y$, so that $x^2 + y^2$ is a nontrivial representation of $0 \bmod p$. But then $x^2 + y^2 = 0$ in $Z/(p)$ gives $-1 = (x/y)^2$, contradicting (i).

(iii) 3.2 shows that $1 < t(2,p) \leqq 3$ (if $-1 = x^2 + y^2$ then $0 = 1^2 + x^2 + y^2$). The rest is then clear from (i) and (ii).

We can now prove 3.1. If $n = p_1^{e_1} \cdots p_l^{e_l}$ (p_i distinct primes, $e_i > 0$), the Chinese Remainder Theorem gives $g(k,n) = \max_i g(k, p_i^{e_i})$ for any k (and $t(k,n) = \min_i t(k, p_i^{e_i})$, which is the last hint for the Exercise of Section 2!). Hence 3.1 is equivalent to the following:

(a) For $p = 2$: $g(2,p) = 1$, $g(2,p^2) = 3$, $g(2,p^e) = 4$ for all $e \geqq 3$.

(b) For $p \equiv 1 \bmod 4$: $g(2,p^e) = 2$ for all $e \geqq 1$.

(c) For $p \equiv 3 \bmod 4$: $g(2,p) = 2$, and $g(2,p^e) = 3$ for all $e \geqq 2$.

To prove (a), check $g(2,2)$, $g(2,4)$ and $g(2,8)$ by hand, and invoke 2.4 (or Lagrange's theorem) to get $g(2,2^e) = 4$ for all $e > 3$.

For (b), plug $g(2,p) = 2$ (from 3.2) and $t(2,p) = 2$ (from 3.3) into 2.2; (c) follows similarly from 3.2, 3.3 and 2.2. Done!

4. The problem mod *p*. As indicated above, the computation of $g(k,n)$ is reduced to the case where n is a prime power by the Chinese Remainder Theorem, and then (almost) to the case where n is prime by the results described in Section 2. Hence in this section we study $g(k,p)$ for primes p.

The even prime, for once, gives no trouble: $g(k,2) = 1$ for all k. *For the remainder of this section, fix the following notation: p is an odd prime, F is $Z/(p)$, G is the (cyclic) group of units of F, $u = p - 1$ is the order of G, $l = (k,u)$ (greatest common divisor), $G^k =$ the subgroup $\{x^k \mid x \in G\}$ of G and ${}_kG =$ the subgroup $\{x \in G \mid x^k = 1\}$ of G.*

4.1 (**[10]**, 7.14) $g(k,p) \leqq l = |G : G^k|$. *In particular, if $p \nmid k$, then $g(k,p^e) \leqq k+1$ for all e.*

Proof: Since $G^k \simeq G/{}_kG$, the equality $l = |G : G^k|$ is eqivalent to $|{}_kG| = l$. Clearly $x^k = 1 \Leftrightarrow x^l = 1$: the order of x divides u, hence divides k iff it divides l. Hence ${}_kG = {}_lG$. Now if σ is a generator for G we have ${}_lG = \{\sigma^{u/l}, \sigma^{2u/l}, \ldots, \sigma^{lu/l} = 1\}$, and $|{}_lG| = l$.

For the inequality $g(k,p) \leqq |G : G^k|$, put $G_i = \{x \in G \mid x$ is a sum i kth powers in $F\}$. Then $G^k = G_1 \subseteq G_2 \subseteq \cdots$. Clearly if $G_i = G_{i+1}$ then $G_i = G_{i+j} \forall j \geqq 1$. Also, for each i,

$\{xy \mid x \in G^k, y \in G_i\} = G_i$, so that if $G_i \subsetneqq G_{i+1}$ then $|G_{i+1}| \geqq |G_i| + |G^k| = |G_i| + u/l$. Hence in the chain $G^k = G_1 \subseteq G_2 \subseteq \cdots \subseteq G$, there are at most $l-1$ strict inclusions, and $G_l = G_{l+j} \forall j \geqq 1$. Thus $G = \bigcup_{i \geqq 1} G_i = G_l$, as claimed. The last statement then follows from 2.2.

REMARK: Part of 4.1 depends only on the fact that G is cyclic, and generalizes: if G is the group of units mod p^e (p an odd prime, $e \geqq 1$) then G is cyclic of order $u = p^e - p^{e-1}$ (see e.g. **[16]**, Theorems 23 and 28) and, as above, $|G: G^k| = l = (k, u)$. However the other half of 4.1, viz. $g(k,p) \leqq |G: G^k|$, uses also the fact that G is the group of non-zero elements of a field, and can fail when p is replaced by p^e, $e > 1$. Mod 27, for example, the only cubes are $0, \pm 1$, and 8 and 10 and their inverses 17 and 19, and $g(3, 3^3) = 4$, whereas $|G: G^3| = (3, 18) = 3$. What goes wrong in the attempt to carry over the proof of 4.1 to the case p^e, $e > 1$, is the assertion: $G_i = G_{i+1} \Rightarrow G_i = G_{i+j} \forall j \geqq 1$. In the example $k = p = e = 3$, in fact, one has $G_1 \subset G_2 = G_3 \subset G_4$, and G_1, G_2, G_4 have 6, 12, 18 elements respectively. We shall return to this phenomenon in Section 7.

4.2 COROLLARY. *$g(k,p) = g(l,p)$ and if $l = 1, 2, u/2$ or u, then $g(k,p) = l$.*

Proof: We saw in proving 4.1 that ${}_kG = {}_lG$. Hence $G^k = G^l$: they are subgroups, of the same order, of the cyclic group G. Certainly, then, $g(k,p) = g(l,p)$. The result for $l = 1$ or 2 then follows from 4.1. Since $\{x^u \mid x \in F\} = \{0, 1\}$, it is clear that $g(u,p) = u$. Finally for $k = u/2$ we have $|G^k| = 2$ from 4.1. But any generator σ for G satisfies $\sigma^k = -1$, since $u = 2k$. Hence $\{x^k \mid x \in F\} = \{0, \pm 1\}$. Thus k itself is not a sum of fewer than k kth powers mod p, so that $g(k,p) = k$.

The preceding results show that to compute $g(k,p)$ for all k it suffices to consider only those k which divide $u = p-1$ and satisfy $3 \leqq k \leqq u/3$. Call such k *relevant* (for p). For relevant k, the bound $g(k,p) \leqq l = k$ can be strengthened considerably:

4.3 THEOREM. *Put $l = (k, u)$ and assume l is not $u/2$ or u. Then $g(k,p) = g(l,p) \leqq [l/2] + 1$ (where $[x]$ = greatest integer in x).*

Proof: This follows from Theorem 2.1, page 11, of **[13]**, which is a corollary of a theorem of Vosper. To apply the cited theorem, observe that $n \geqq [l/2] + 1$ implies $(2n-1)((p-1)l) + 1 \geqq p$.

4.4. COROLLARY.

$$g(3,p) = \begin{cases} 3 & \text{if} \quad p = 7 \\ 2 & \text{if} \quad p \equiv 1 \bmod 3,\ p \neq 7 \\ 1 & \text{for all other primes } p. \end{cases}$$

Proof: If $p \not\equiv 1 \bmod 3$ then $(3, p-1) = 1$ and $g(3,p) = 1$ by 4.1. $g(3,7) = 3$ by 4.2. For $p \equiv 1 \bmod 3$, $p \neq 7$, we have $g(3,p) \leqq [3/2] + 1 = 2$ by 4.3, and clearly $g(3,p) > 1$ (cf. 4.1).

REMARK: For all $p \leqq 29$ and all relevant k, one actually has $g(k,p) = [k/2] + 1$. For $p = 31$, the relevant k are 3, 5, 6, 10. For $k = 3$, $g(k, 31) \leqq [k/2] + 1$ is an equality by 4.4. For $k = 5$ and $k = 6$ one finds $g(k, 31) = [k/2] + 1$ by direct computation. For $k = 10$, one has $G^{10} = \{1, 5, 25\}$, and $g(10, 31) = 5 < 6 = [10/2] + 1$. This is the only example with $p < 37$ in which the inequality of 4.3 is strict, although it is "clear" that the bound $[k/2] + 1$ becomes worse as p gets large. Precise computation of $g(k,p)$ for relevant $k > 3$ seems to be a difficult problem. There are asymptotic results; for example in **[5]** it is shown that $\gamma(k)$, the maximum of $g(k,p)$ over all p for which k is relevant, behaves like $\sqrt{k}$ for large k.

In a similar vein, the general theory of diagonal equations over finite fields can be applied to improve the bound $g(k,p) \leqq [k/2] + 1$ of 4.3 for relevant k, *provided p is large compared to k*. For example, fix $r \geqq 1$ and $0 \neq b \in \boldsymbol{Z}/(p)$ and let $N(r,b)$ be the number of solutions (in $(\boldsymbol{Z}/(p))^r$) to $X_1^k + \cdots + X_r^k = b$, where k is relevant for p. Then it follows from Cor. 1 on p. 57 of **[11]** that

$$|N(b,r) - p^{r-1}| \leqq (k-1)^r p^{(r-1)/2}$$

(Note that this estimate is independent of b.) In particular, $N(b, r) \geqq p^{r-1} - (k-1)^r p^{(r-1)/2}$, so that $N(b, r) > 0$ provided $p^{r-1} > (k-1)^r p^{(r-1)/2}$, or in other words, $p > (k-1)^{2r/(r-1)}$. Since $g(k, p)$ is clearly the smallest r for which $N(r, b) > 0$ for all b, we conclude that for relevant k,

$$g(k,p) = 2 \quad \text{if} \quad p > (k-1)^4, \qquad \leqq 3 \quad \text{if} \quad p > (k-1)^3 \cdots$$
$$\leqq r \quad \text{if} \quad p > (k-1)^{2r/(r-1)} \cdots \leqq [k/2] \quad \text{if} \quad p > (k-1)^{2[k/2]/([k/2]-1)},$$
$$\leqq [k/2] + 1 \text{ in any case (by 4.3).}$$

Exercise: Verify the following results for everybody's favorite small prime:

$$g(k, 17) = \begin{cases} 1 & \text{if} \quad k \text{ is odd} \\ 2 & \text{if} \quad k \equiv \pm 2 \text{ or } \pm 6 \bmod 16 \\ 3 & \text{if} \quad k \equiv \pm 4 \bmod 16 \\ 8 & \text{if} \quad k \equiv 8 \bmod 16 \\ 16 & \text{if} \quad k \equiv 0 \bmod 16. \end{cases}$$

REMARK: If we try to extend the foregoing results, on sums of kth powers in the prime fields $\boldsymbol{Z}/(p)$, to arbitrary finite fields, there is one immediate obstacle: given a finite field F and an integer k, it need not be true that every element of F is a sum of kth powers. (For example, the only cubes in the field with 4 elements are 0 and 1.) These exceptional cases can be classified (at least when k is prime; see **[20]**, Theorem 5); and it is known that for any finite field F (with, say, q elements) and any k, every element of F which is a sum of kth powers is a sum of $(k, q-1)$ kth powers (*ibid.*, Theorem 1; see also **[10]**, 7.14i). 4.3 gives better bounds for the number of kth powers needed, in the special case when F is a prime field, but Vosper's theorem, upon which 4.3 depends, works only in additive groups of prime order. There are more general versions, for arbitrary abelian groups (see **[12]** and its bibliography), which might conceivably lead to better bounds in the general case in the same way that Vosper's theorem led to 4.3.

5. Sums of cubes mod n. In this section we prove the following theorem, the analogue for cubes of 3.1.

5.1. *Let $r \geqq 1$ and $n > 1$ be integers. Then every element of $\boldsymbol{Z}/(n)$ is a sum of r cubes if and only if one of the following holds*:

(1) $r = 1$ *and*: $9 \nmid n$, $p^2 \nmid n$ *for all primes* $p \equiv 2 \bmod 3$, $p \nmid n$ *for all primes* $p \equiv 1 \bmod 3$;
(2) $r = 2$ *and* $9 \nmid n$, $7 \nmid n$;
(3) $r = 3$ *and* $9 \nmid n$;
(4) $r \geqq 4$ (*and no condition on* n).

As before, the Chinese Remainder Theorem makes it clear that 5.1 is equivalent to:

$$g(3, p^e) = \begin{cases} 3 & \text{if} \quad p = 7,\ e \geqq 1 \\ 2 & \text{if} \quad p \equiv 1 \bmod 3,\ p \neq 7,\ e \geqq 1 \\ 2 & \text{if} \quad p \equiv 2 \bmod 3,\ e > 1 \\ 1 & \text{if} \quad p \equiv 2 \bmod 3,\ e = 1 \\ 1 & \text{if} \quad p = 3,\ e = 1 \\ 4 & \text{if} \quad p = 3,\ e > 1. \end{cases}$$

For $p = 7$, $g(3, p) = 3$ by 4.4 and $g(3, p^e) = 3$ for all $e \geq 1$ by 2.2. For $p \equiv 1 \bmod 3$, $p \neq 7$, we have

$g(3,p)=2$ by 4.4 and $g(3,p^e)=2$ for all $e \geqq 1$ by 2.2. For $p \equiv 2 \bmod 3$, $g(3,p)=1$ by 4.4 and $g(3,p^e) \leqq 2$ for all $e>1$ by 2.2; on the other hand $g(3,p^e) \geqq 2$ by:

5.2 LEMMA. *For all $k>1$ and $e>1$, p isn't a k-th power* $\bmod p^e$; *consequently $g(k,p^e) \geqq 2$.*

Proof: If $p=x^k+mp^e$ then $p \mid x$. But then $p^2 \mid x^k$, which implies $p^2 \mid p$, a contradiction.

To complete the proof of 5.1 we have to compute $g(3,3^e)$ for all $e \geqq 1$. One verifies easily that $g(3,3)=1$ and $g(3,9)=g(3,27)=4$. Now apply 2.4 (with $k,p,c=3,3,1$ respectively) to get $g(3,3^e)=4$ for all $e>1$. Done!

6. Examples. It is easy to assemble a general result for $g(k,n)$ from what has preceded:

6.1. *Let $n=p_1^{e_1}\cdots p_l^{e_l}$ (p_i distinct primes, $e_i>0$) and for each i define $c_i \geqq 0$ by: $p_i^{c_i} \| k$. Then*

$$\max_i g(k,p_i) \leqq \max_i g(k,p_i^{e_i}) = g(k,n) \leqq \max_i \left(g(k,p_i^{2c_i+1})+\varepsilon_i\right)$$

where $\varepsilon_i=0$ if $c_i>0$, or if $e_i \leqq 2c_i+1$, or if k is odd and $(k,p_i-1)>1$, and $\varepsilon_i=1$ in general.

Proof: $g(k,n)=\max_i g(k,p_i^{e_i})$ by the Chinese Remainder Theorem, and the first inequality is clear. For the rest, it suffices to show $g(k,p_i^{e_i}) \leqq g(k,p_i^{2c_i+1})$ if $c_i>0$, $g(k,p_i^{e_i})=g(k,p_i)$ if k is odd and $(k,p_i-1)>1$, and $g(k,p_i^{e_i}) \leqq 1+g(k,p_i)$ whenever $c_i=0$. The first of these follows from 2.4. For the others, use 2.2 and note that k odd gives $t(k,p_i)=2$ while $(k,p_i-1)>1$ gives $g(k,p_i) \geqq 2$ (4.1).

6.2 COROLLARY. *Assume $(k,n)=1$; then*

$$\max g(k,p) \leqq g(k,n) \leqq 1+\max g(k,p)$$

(the max *is over all prime divisors p of n), with equality on the left if n is squarefree.*

Proof: The last statement is just the Chinese Remainder Theorem since n is squarefree iff all $e_i=1$. The rest follows from 6.1 since $(k,n)=1$ iff all $c_i=0$.

Call a prime divisor p of n *corelevant* (for k) if $(p-1,k)>1$ (equivalently, by 4.1, if $g(k,p)>1$). (Note that "p is corelevant for k" is not the same as "k is relevant for p".)

6.3 COROLLARY. *Assume $(k,n)=1$; then*:

CASE 1: *n has no corelevant prime divisors. Then $g(k,n)=1$ if n is squarefree, and 2 otherwise.*

CASE 2: *n has corelevant prime divisors. Then if k is odd, or if n is squarefree, $g(k,n)=\max g(k,p)$; and in any case $g(k,n)=\max g(k,p)$ or $1+\max g(k,p)$; where each max is over all corelevant prime divisors p.*

Proof: In Case 1, 6.2 yields $1 \leqq g(k,n) \leqq 2$. When n is not squarefree, therefore, $g(k,n)=2$ by 5.2. Clearly when there are corelevant primes, $\max g(k,p)$ over all such p is the same as $\max g(k,p)$ over *all* prime divisors p. Finally, the special conclusion when k is odd follows from 6.1.

6.3 computes $g(k,n)$ explicitly, at least when $(k,n)=1$ and n is squarefree or k is odd, modulo computation of $g(k,p)$ for the corelevant prime divisors p of n. For these $g(k,p)$ we have in general only 4.2 and 4.3. However, if n has no corelevant prime divisors $\geqq 37$, this gives a precise computation (cf. Remark following 4.4). To illustrate, here are two examples:

We compute $g(7,n)$ for $n=$

$$2^{e_1}\cdot 3^{e_2}\cdot 5^{e_3}\cdot 11^{e_4}\cdot 13^{e_5}\cdot 17^{e_6}\cdot 19^{e_7}\cdot 23^{e_8}\cdot 29^{e_9}\cdot 31^{e_{10}}\cdot 37^{e_{11}}\cdot 41^{e_{12}}\cdot 47^{e_{13}}\cdot 53^{e_{14}}\cdot 101^{e_{15}}$$

($e_i \geqq 0$ arbitrary, not all 0). The only corelevant prime is 29, so $g(7,n)=g(7,29)=4$ if $e_9>0$. If $e_9=0$ then $g(7,n)=1$ if all $e_i \leqq 1$, and $g(7,n)=2$ otherwise, by Case 1 of 6.3.

We compute $g(33,n)$ for $n=$

$$2^{e_1}\cdot 5^{e_2}\cdot 7^{e_3}\cdot 13^{e_4}\cdot 17^{e_5}\cdot 19^{e_6}\cdot 23^{e_7}\cdot 29^{e_8}\cdot 31^{e_9}\cdot 41^{e_{10}}\cdot 53^{e_{11}}\cdot 599^{e_{12}}$$

($e_i \geqq 0$ arbitrary, not all 0). The corelevant prime divisors are 7, 13, 19, 23 and 31, and if these all actually occur, $g(33,n)$ is the max of $g(33,7)=g(3,7)=3$, $g(33,13)=g(3,13)=2$, $g(33,19)=g(3,19)=2$, $g(33,23)=g(11,23)=11$, and $g(33,31)=g(3,31)=2$. Thus $g(33,n)=11$ if $e_7>0$, $g(33,n)=3$ if $e_7=0$ but $e_3>0$, $g(33,n)=2$ if $e_7=e_3=0$ but one of e_4, e_6, e_9 is >0, or if $e_7=e_3=e_4=e_6=e_9=0$ and some e_i is $\geqq 2$, and $g(33,n)=1$ in the remaining case, when $e_7=e_3=e_4=e_6=e_9=0$ and all $e_i \leqq 1$.

7. Local-global questions. Any computation of $g(k,n)$ yields lower bounds for $g(k)$ and $G(k)$ (this is spelled out below), and there is accordingly some hope that our work mod n may have consequences for the "global" Waring problem (in $\boldsymbol{Z}$). The aim of this section is to explore that theme. Unfortunately, we are able to show that none of the obvious applications gives anything new for the global problem, although in one case (cf. 7.3, 7.4, 7.9 below) this negative result is non-trivial and depends on a substantial theorem (7.10). The fact that we get no new results on the global problem is hardly surprising, since, as we indicated in the Introduction, our methods in this paper are very weak in comparison with the machinery typically brought to bear on Waring's problem.

Clearly $g(k,n)\leqq g(k)$ always. Define $g'(k)=\sup_n g(k,n)$; then $g'(k)\leqq g(k)$. The conjecture $g'=g$ is quickly defeated: we found in Section 5 that $g'(3)=4$, whereas $g(3)$ is known to be 9 (and in fact 23 is certainly not the sum of fewer than 9 cubes). Since, as we noted above, $g(k)$ is completely known except for $k=4$, the interest of the inequality $g'\leqq g$ is that one might be able to improve the lower bound $\bar{g}(4)=19\leqq g(4)$ by computing $g'(4)$. Our experience for cubes suggests that such a hope is doomed, as the following confirms:

7.1 *Let* $k=p_1^{e_1}\cdots p_l^{e_l}$ *(p_i distinct primes, $e_i>0$). Then* $g'(k)\leqq \max(k+1, \max_i g(k,p_i^{2e_i+1}))$. *If* $g(k,p_i^{2e_i+1})\geqq k+1$ *for some* i, *then* $g'(k)=\max_i g(k,p_i^{2e_i+1})$. *In particular,* $g'(4)=15$.

Proof: By the Chinese Remainder Theorem, $g'(k)$ is the sup, over all primes q and all $e\geqq 1$, of $g(k,q^e)$. If q is not one of the p_i then $g(k,q^e)\leqq k+1$ by 4.1; if $q=p_i$ then $g(k,q^e)\leqq g(k,p_i^{2e_i+1})$ by 2.4. This proves everything except $g'(4)=15$, and shows that $g'(4)\leqq \max(5, g(4,32))$. The 4th powers mod 32 are 0, 1, 16 and 17, and clearly $g'(4)=g(4,32)=15$.

The equation $g'(4)=15$ says that 15 is the smallest s with the property that in every arithmetic progression there is a number which is the sum of s 4th powers. In fact, 15 is also the smallest s with the property that in every arithmetic progression there are *infinitely many* numbers which are sums of s 4th powers. See **[8b]**, Theorem 1.

Note that thanks to 7.1, we do not need Hilbert's theorem ($g(k)<\infty$) to know that $g'(k)$ is well-defined! Also, 7.1, together with Euler's lower bound $\bar{g}(k)$ for $g(k)$ (see Section 1), suggests that $g(k)-g'(k)$ is "usually" large.

Let $G(k)$ denote the smallest r such that every positive integer, *except for possibly finitely many*, is a sum of r kth powers of positive integers. Then clearly $g'(k)\leqq G(k)\leqq g(k)$. G studies "essentially" the same phenomena as g, but it ignores peculiarities due to small numbers. While $g(k)$ can be computed explicitly for all $k\neq 4$, very little is known about G. For example, it is easy to see that $G(k)\geqq k+1$ for all k (see **[9]**, Theorem 394), but there is no odd prime p for which one can at present prove either $G(p)=p+1$ or $G(p)>p+1$. Even $G(3)$ is unknown: one has $G(3)\geqq 4$ (indeed, we have seen that $g(3,9)=4=g'(3)\leqq G(3)$); it is known that 23 and 239 are the only integers requiring nine cubes **[4]**, so that $G(3)\leqq 8$; and in fact $G(3)\leqq 7$ (see **[21]**). But $4\leqq G(3)\leqq 7$ is all that's known about $G(3)$. For $p=2$ we have $g'(2)\geqq g(2,8)=4$, and $4\leqq g'(2)\leqq G(2)\leqq g(2)=4$ gives $G(2)=4$.

The reasonable conjecture that $g'=G$ in general is quickly defeated: we just found $g'(4)=15$, whereas an easy argument shows $G(4)\geqq 16$ (**[9]**, Theorem 395).* The inequality $g'\leqq G$ does,

* And a less easy argument shows $G(4)=16$, **[3]**. $G(2)=4$ and $G(4)=16$ are in fact the *only* values of k for which $G(k)$ is known exactly.

however, leave open the possibility that one might be able to improve known lower bounds for G by computing g'. Explicitly:

7.2 *Let* $k = p_1^{e_1} \cdots p_l^{e_l}$ (p_i *distinct primes,* $e_i > 0$). *Consider the following three conditions (which may or may not hold):*

(1) $g'(k) > k+1$.

(2) $g(k, p_i^{2e_i+1}) > k+1$ *for some* i.

(3) $G(k) > k+1$.

Then (1) *and* (2) *are equivalent, and they imply* (3).

Proof: The equivalence of (1) and (2) is clear from 7.1, and (1) $\Rightarrow$ (3) is clear since $g'(k) \leqq G(k)$.

If we want to use 7.2 to improve known lower bounds for $G(k)$ by computing $g'(k)$, the most important case will perhaps be when k is prime. This is because for odd primes p, $p+1$ is the best known lower bound for $G(p)$, whereas for some composite k, better lower bounds are known for $G(k)$ (see **[9]**, Theorem 400). When k is prime, 7.2 becomes:

7.3 *If* p *is prime then* $g'(p) > p+1 \Leftrightarrow g(p, p^3) > p+1$, *in which case* $G(p) > p+1$.

The point in 7.3 is that $g(p, n) > p+1$ for *some* n if and only if $g(p, n) > p+1$ for $n = p^3$. Because of this and its relevance to $G(p)$, one is interested in computing $g(p, p^3)$. The following surprising result simplifies that task:

7.4 *If* p *is an odd prime,* $g(p, p^2) = g(p, p^3)$.

Before proving 7.4 we introduce some notation and prove some lemmas (7.5–7.8).

For $x \in \mathbf{Z}$ let x' (resp. x'') denote the image of x in $\mathbf{Z}/(p^2)$ (resp. $\mathbf{Z}/(p^3)$). Let f: $\mathbf{Z}/(p^3) \to \mathbf{Z}/(p^2)$ be the natural map, so that $x' = f(x'')$ for $x \in \mathbf{Z}$. Write $U(n)$ for the group of units of $\mathbf{Z}/(n)$, and $U(n)^m$ for its subgroup $\{x^m \mid x \in U(n)\}$. Note that $f(x) \in U(p^2) \Leftrightarrow x \in U(p^3)$.

7.5 *Let* p *be a prime and* $1 \leqq e \leqq p$. *Then in* $\mathbf{Z}/(p^e)$, *every non-zero* p*-th power is a unit.*

Proof: Suppose $x \in \mathbf{Z}$ is such that $\bar{x}$ is not a unit in $\mathbf{Z}/(p^e)$ (where $\bar{x}$ denotes the image of x $\bmod p^e$). Then $p \mid x$, so that $p^p \mid x^p$, and since $e \leqq p$, $p^e \mid x^p$ too. Thus $\bar{x}^p = 0$.

We will use 7.5 for odd p and $e = 2$ or 3, in which case it says that $U(p^e)^p$ is the set of non-zero pth powers $\bmod p^e$.

We have already seen that $U(p^2)^p$ and $U(p^3)^p$ have $p-1$ and p^2-p elements, respectively. The following result gives a convenient way of listing them explicitly:

7.6 *Let* p *be an odd prime. Then* $U(p^2)^p = \{x'^p \mid 1 \leqq x \leqq p-1\}$ *and* $U(p^3)^p = \{(x^p + lp^2)'' \mid 1 \leqq x \leqq p-1, 0 \leqq l \leqq p-1\}$.

Proof: Clearly for any $x \in \mathbf{Z}$ we have $(x+p)^p \equiv x^p \bmod p^2$, so we get all the non-zero pth powers $\bmod p^2$ by taking the x'^p for $1 \leqq x \leqq p-1$. (A similar argument shows $U(p^3)^p = \{x''^p \mid 1 \leqq x \leqq p^2-1, p \nmid x\}$, since $(x+p^2)^p \equiv x^p \bmod p^3$ $\forall x \in \mathbf{Z}$, but we have to prove more.)

We show first that if x, y, z are integers, not divisible by p, such that $x \equiv y^p \bmod p^3$ and $x \equiv z \bmod p^2$, then $z = (y+tp)^p \bmod p^3$ for some t (in fact for exactly one t satisfying $0 \leqq t \leqq p-1$). In particular this shows that if $x \equiv z \bmod p^2$ and $x'' \in U(p^3)^p$ then $z'' \in U(p^3)^p$; in other words, if we have an element x'' of $U(p^3)^p$, we can get $p-1$ others by forming $(x+lp^2)''$, $1 \leqq l \leqq p-1$, for clearly the p integers $\{x + lp^2 \mid 0 \leqq l \leqq p-1\}$ are distinct $\bmod p^3$. To prove the claim, choose $a \in \mathbf{Z}$ such that $x = z + ap^2$, and observe that $\bmod p^3$ we have

$$(y+tp)^p \equiv y^p + p \cdot y^{p-1} \cdot tp \equiv x + ty^{p-1}p^2 \equiv z + (a + ty^{p-1})p^2,$$

since $\binom{p}{2} = p(p-1)/2$ is divisible by p. Since y^{p-1} is a unit $\bmod p^3$, we can choose t to get $a + ty^{p-1} \equiv 0 \bmod p^3$, and with this choice we have $(y+tp)^p \equiv z \bmod p^3$ as desired.

To finish the proof, note that if x, y are distinct integers, $1 \leqq x, y \leqq p-1$, then $x^p \not\equiv y^p \bmod p^3$. (In fact, since $x^p \equiv x$ and $y^p \equiv y \bmod p$, we even have $x^p \not\equiv y^p \bmod p$.) Thus $\{x''^p \mid 1 \leqq x \leqq p-1\}$ is a set of $p-1$ distinct elements of $U(p^3)^p$. Translating each of them by multiples of p^2 as in the previous paragraph, we get $(p-1)p$ elements (they are clearly distinct), i.e. *all* the elements. This completes the proof of 7.6.

Everything we need to prove 7.4 follows from 7.6:

7.7 *Let p be an odd prime. Then*:
(i) $f(x) \in U(p^2)^p \Leftrightarrow x \in U(p^3)^p$, *and*
(ii) *If $f(x) = 0$ then x is a sum of two p-th powers in $\mathbf{Z}/(p^3)$.*

Proof: (i) is clear form 7.6: each element of $U(p^2)^p$ visibly has p pre-images in $U(p^3)^p$; but it has only p pre-images altogether! For (ii), just note that ± 1 are pth powers; it is then clear from 7.6 that anything of the form $lp^2 \pm 1$ is a pth power, and in particular, that any multiple of p^2 is a sum of two pth powers, $\bmod p^3$.

7.8 *Let p be an odd prime, suppose $x \in \mathbf{Z}/(p^3)$ is such that $f(x) \neq 0$, and assume that $f(x)$ is a sum of r p-th powers in $\mathbf{Z}/(p^2)$ (for some $r \geqq 1$). Then x is a sum of r p-th powers in $\mathbf{Z}/(p^3)$.*

Proof: First consider the case $r = 1$: $f(x)$ is a non-zero pth power in $\mathbf{Z}/(p^2)$, and we want to show x is a pth power in $\mathbf{Z}/(p^3)$. Since non-zero pth powers in $\mathbf{Z}/(p^2)$ or $\mathbf{Z}/(p^3)$ are units (7.5), it suffices to show

$$\{x \in U(p^3) \mid f(x) \in U(p^2)^p\} \subseteq U(p^3)^p.$$

But this is part of 7.7.

Now suppose $r > 1$. We have $f(x) = y_1^p + y_2^p + \cdots + y_r^p$ in $\mathbf{Z}/(p^2)$. We may assume (renumbering if necessary) that $y_1^p \neq 0$. Choose x_i in $\mathbf{Z}/(p^3)$ so that $f(x_i) = y_i$, $2 \leqq i \leqq r$, and put $z = x - (x_2^p + \cdots + x_r^p) \in \mathbf{Z}/(p^3)$. Then $f(z)$ is a non-zero pth power in $\mathbf{Z}/(p^2)$. By the case $r = 1$, z is a pth power in $\mathbf{Z}/(p^3)$; done.

We can now prove 7.4, which says $g(p, p^2) = g(p, p^3)$ for odd primes p. We have to show $g(p, p^3) \leqq g(p, p^2)$, and since $g(p, p^3) \geqq 2$, it suffices to show that if $x \in \mathbf{Z}/(p^3)$ and $f(x)$ is a sum of r pth powers in $\mathbf{Z}/(p^2)$ (for some $r \geqq 1$) then x is a sum of $\max(r, 2)$ pth powers in $\mathbf{Z}/(p^3)$. If $f(x) \neq 0$, this follows from 7.8, and if $f(x) = 0$ it follows from 7.7. (ii).

REMARK: Once we had 7.7, it was easy to prove 7.4. We derived 7.7 from 7.6, which describes $U(p^2)^p$ and $U(p^3)^p$ very concretely. Alternatively, one can bypass 7.6 and give direct proofs for 7.7, as follows. First, show that if $f: G \twoheadrightarrow H$ is a surjective homomorphism of finite cyclic groups and G_1 (resp. H_1) is a subgroup of G (resp. H) such that $[G: G_1] = [H: H_1]$, then $f^{-1}(H_1) = G_1$. (This is a simple exercise.) Apply to $G, H, G_1, H_1 = U(p^3), U(p^2), U(p^3)^p, U(p^2)^p$ respectively, to get 7.7. (i). For 7.7. (ii) it is enough to show that any integer of the form $lp^2 \pm 1$ is a pth power $\bmod p^3$, and in fact the binomial theorem gives $(\pm 1 + lp)^p \equiv (\pm 1)^p + p(\pm 1)^{p-1} lp = lp^2 \pm 1 \bmod p^3$ since $\binom{p}{2} = p(p-1)/2$ is divisible by p.

Exercise: Let p be an odd prime. Show that $g'(p) = g(p, p^2)$ if $g(p, p^2) \geqq p+1$; $g'(p) = p$ if $g(p, p^2) < p+1$ and $2p+1$ is prime; and $1 < g'(p) \leqq p$ in the remaining case. This shows, for example, that $g'(5) = 5$; since $G(k) \geqq k+1$ for all k, we conclude that an inequality $g'(k) < G(k)$, noted above for $k = 4$, is possible even when k is prime. It would be interesting to know how $G(k) - g'(k)$ behaves, as a function of k.

With the preceding results at hand, we have that $g(p, n) > p+1$ for *some* n if, and only if, $g(p, p^2) > p+1$. Thus, to use our work mod n to improve the lower bound $G(p) \geqq p+1$, we seek odd primes p satisfying $g(p, p^2) > p+1$. The possibility of such an application is certainly tantalizing, but the following shows it is an idle dream:

7.9 *For all odd primes p,* $g(p,p^3)=g(p,p^2)\leqq p+1$.

For the proof of 7.9, we need a lemma:

7.10 *Let A be a subset of* $\boldsymbol{Z}/(n)$ *such that* $0\in A$ *and* $A-\{0\}\subseteq U(n)$. *Let B be any subset of A and write* $A+B=\{a+b\mid a\in A, b\in B\}$. *Then either* $A+B=\boldsymbol{Z}/(n)$ *or* $|A+B|\geqq|A|+|B|-1$, *where* $|X|$ *denotes the number of elements in X.*

7.10 is a theorem of Chowla; see **[13]**, Cor. 1.2.4 on page 3. We will use it in the following form:

7.11 *Let A be as in* 7.10 *and assume* $r\geqq(n-1)/(|A|-1)$. *Then* $rA=\boldsymbol{Z}/(n)$, *where, for any integer* $k\geqq 1$, kA *means* $\{a_1+\cdots+a_k\mid a_i\in A, 1\leqq i\leqq k\}$.

Proof: If $rA\subsetneqq\boldsymbol{Z}/(n)$ then clearly $kA\subsetneqq\boldsymbol{Z}/(n)$ for all k, $2\leqq k\leqq r$. Applying 7.10 to $2A\subsetneqq\boldsymbol{Z}/(n)$ (with $B=A$) gives $|2A|\geqq 2|A|-1$. Applying 7.10 to $3A\subsetneqq\boldsymbol{Z}/(n)$ (with $B=2A$) gives $|3A|\geqq|2A|+|A|-1\geqq 3|A|-2$. Continuing, we find $|rA|\geqq r|A|-(r-1)$. But

$$r\geqq\frac{n-1}{|A|-1}\quad\text{implies}\quad r|A|-(r-1)\geqq n,$$

and this contradiction completes the proof.

It is convenient to isolate a special case of 7.11:

7.12 *Assume k and n are such that every non-zero k-th power in* $\boldsymbol{Z}/(n)$ *is a unit. Then*

$$g(k,n)\leqq\left\{\frac{n-1}{|U(n)^k|}\right\},$$

where $\{x\}$ *denotes the smallest integer l such that* $x\leqq l$, *i.e.,* $\{x\}=x$ *if x is an integer and* $[x]+1$ *otherwise.*

Proof: Apply 7.11 with $A=\{x^k\mid x\in\boldsymbol{Z}/(n)\}$; note that

$$r\geqq\frac{n-1}{|A|-1}=\frac{n-1}{|U(n)^k|}\quad\text{if, and only if,}\quad r\geqq\left\{\frac{n-1}{|U(n)^k|}\right\}.$$

(For example, when n is prime, 7.12 gives $g(k,n)\leqq(k,n-1)$, which we already knew.)

The proof of 7.9 is now clear: apply 7.12 with $n=p^2$, $k=p$, and observe that

$$\left\{\frac{p^2-1}{|U(p^2)^p|}\right\}=\left\{\frac{p^2-1}{p-1}\right\}=p+1.$$

Note that 7.4, which gave $g(p,p^3)=g(p,p^2)$, is essential in the foregoing arguments; if we use 7.12 directly to estimate $g(p,p^3)$ we find

$$g(p,p^3)\leqq\left\{\frac{p^3-1}{p^2-p}\right\}=\left\{\frac{p^2+p+1}{p}\right\}=\left\{p+1+\frac{1}{p}\right\}=p+2!$$

Note that 7.9 is trivially "best possible": $g(2,4)=3$ and $g(3,9)=4$; this shows also that 7.10 is best possible, for if we had $|A+B|\geqq|A|+|B|$, the arguments above would give $g(p,p^2)\leqq p$.

It is tempting to try to prove 7.9 directly, circumventing Chowla's theorem and its relatively difficult proof, by an argument like that used to prove 4.1. Namely, let $A=\{x^p\mid x\in\boldsymbol{Z}/(p^2)\}$; we want to show $(p+1)A=\boldsymbol{Z}/(p^2)$. Observe that $A\subseteq 2A\subseteq 3A\subseteq\ldots$, $\bigcup_{n\geqq1}nA=\boldsymbol{Z}/(p^2)$, and if $nA=(n+1)A$ then $nA=(n+j)A$ for all $j\geqq1$. Now $p-1$ is the number of units in A; suppose we knew

$$(*)\qquad\text{if}\quad nA\subsetneqq(n+1)A\quad\text{then}\quad|(n+1)A|\geqq|nA|+p-1.$$

We could then conclude, as in the proof of 4.1, and as in the proof of 7.11, that $\boldsymbol{Z}/(p^2)=\bigcup_{n\geqq1}nA=(p+1)A$. There seems to be no obvious reason for (*) to be true, though it would be surprising if it

were false. If one knew that whenever $nA\subsetneqq(n+1)A$ then necessarily $U(p^2)\cap nA\subsetneqq U(p^2)\cap(n+1)A$, one could conclude (*) immediately and be done. However, it is just not true that whenever one has a sum of $n+1$ pth powers in $Z/(p^2)$ which is not a sum of n pth powers, then one has a *unit* with those properties. For example in $Z(9)$ we have $A\subsetneqq 2A\subsetneqq 3A\subsetneqq 4A = Z/(9)$, but if we put $G_n = nA \cap U(p^2)$, we find the picture $G_1\subsetneqq G_2 = G_3\subsetneqq G_4$, as in the example preceding 4.2.

Some interesting (and, no doubt, difficult) questions remain:

(1) Is the converse, "$G(p)>p+1$ implies $g(p,p^2)>p+1$", valid? In other words, if there are infinitely many positive integers which cannot be written as a sum of $p+1$ or fewer positive pth powers, is there necessarily an arithmetic progression $a+np^2$, $n\in Z$, of integers, none of which can be written as a sum of $p+1$ pth powers in Z? For any prime p for which this is true, 7.9 implies $G(p)=p+1$. (Compare the exercise above 7.9, where we saw that $x_1^5+\cdots+x_5^5$ misses infinitely many values, but does *not* miss any arithmetic progression.)

(2) Although we have seen that the attempt to improve the lower bound $G(p)\geqq p+1$ by computing $g'(p)$ is doomed, there is still the possibility that computations of $g'(k)$ for *composite* k may improve existing lower bounds for $G(k)$. 7.1 shows that this will involve finding good lower bounds for $g(k,p^{2e+1})$ where $p^e\,\|\,k$. 7.12 provides an upper bound:

7.13 *Let* $k=2^e p_1^{e_1}\cdots p_i^{e_i}$ (p_i *distinct odd primes*, $e\geqq 0$, $e_i>0$). *Fix* i, *and for each* $j\neq i$ *define* $a_j\geqq 0$ *by* $p_j^{a_j}\,\|\,(p_i-1)$, *and define* f *by* $2^f\,\|\,(p_i-1)$. *Then* $g(k,p_i^{2e_i+1})\leqq\{abc/p_i^{e_i}\}$, *where* $a=\sum_{\nu=0}^{2e_i}p_i^{\nu}$, $b=\prod_{j\neq i}p_j^{\min(a_j,e_j)}$, *and* $c=2^{\min(e,f)}$.

We leave to the reader the pleasant task of deriving 7.13 from 7.12. By way of example, let us investigate $g(p^e,p^{2e+1})$ for an odd prime p. There are no a_j's, and we find

$$\begin{aligned} g(p^e,p^{2e+1}) &\leqq \left\{\frac{p^{2e}+p^{2e-1}+\cdots+p+1}{p^e}\right\}\\ &=\left\{p^e+p^{e-1}+\cdots+p+1+\left(\frac{1}{p}+\cdots+\frac{1}{p^e}\right)\right\}=1+\sum_{\nu=0}^{e}p^{\nu}\\ &=1+\frac{p^{e+1}-1}{p-1}.\end{aligned}$$

To complete 7.13 we should compute $g(k,2^{2e+1})$ where $2^e\,\|\,k$. If k is odd ($e=0$) or $k=2$ or $k=4$, the results are trivial: $g(k,2)=1$, $g(2,8)=4$, $g(4,32)=15$. For even $k>4$, 7.12 applies (this is easily checked; it boils down to the observation that if $4<k=2^ek'$ with k' odd, then $2e+1\leqq k$). Using the fact ([**16**], page 70) that

$$U(2^{2e+1})\cong C_2\times C_{2^{2e-1}} \text{ for } e\geqq 1,$$

where C_n denotes a cyclic group of order n, we have

$$|\,U(2^{2e+1})^k\,| = 2^{2e-1}/(k,2^{2e-1}) = 2^{2e-1}/2^e = 2^{e-1},$$

so that $g(k,2^{2e+1})\leqq\{(2^{2e+1}-1)/2^{e-1}\}=7$ if $e=1$, 2^{e+2} if $e>1$.

References

1. R. G. Archibald, Waring's problem: squares, Scripta Math., 7 (1940) 33–48.

2. J. Chen, Waring's problem for $g(5)=37$, Sci. Sinica, 13 (1964) 335 and 1547–1568; = Chinese Math. Acta, 6 (1965) 105–127.

3. H. Davenport, On Waring's problem for fourth powers, Ann. of Math, 40 (1939) 731–747.

3a. ———, Analytic Methods for Diophantine Equations and Diophantine Inequalities, Ann Arbor Publishers, 1962.

4. L. E. Dickson, All integers except 23 and 239 are sums of eight cubes, Bull. A.M.S., 45 (1939) 588–591.

5. M. M. Dodson and A. Tietäväinen, A note on Waring's problem in $GF(p)$ (to appear in Acta Arith.).

6. W. J. Ellison, Waring's problem, this MONTHLY, 78 (1971) 10–36.

7. M. Griffin and M. Krusemeyer, Matrices as sums of squares, to appear in Lin.&Multilin. Alg.

8a. G H. Hardy and J. E. Littlewood, Some Problems of 'Partitio Numerorum' IV, Math. Z., 12 (1922) 161–188.

8b. ———, Some Problems of 'Partitio Numerorum' VIII: the Number $T(k)$ in Waring's Problem, Proc. London Math. Soc., 28 (1927) 518–542.

9. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, Oxford University Press, 4th edition (1960), reprinted with corrections 1968.

10. J.-R. Joly, Sommes de puissances d-ièmes dans un anneau commutatif, Acta Arith., 17 (1970) 37–114.

11. ———, Equations et variétés algébriques sur un corps fini, l'Enseignement Math., 19 (1973) 1–117.

12. J. H. B. Kemperman, On small sumsets in an abelian group, Acta Math., 103 (1960) 63–88.

12a. K. Mahler, On the fractional parts of the powers of a rational number, II, Mathematika, 4 (1957) 122–124.

13. H. B. Mann, Addition Theorems, Interscience, New York, 1965.

14. D. J. Newman, A Simplified Proof of Waring's Conjecture, Mich. Math. J., 7 (1960) 291–295.

15. H.-H. Ostmann, Additive Zahlentheorie (two volumes), Springer-Ergebnisse, Berlin, 1956.

16. F. Richman, Number Theory: An Introduction to Algebra, Brooks-Cole (Wadsworth), Belmont, Calif., 1971.

17. J.-P. Serre, A Course in Arithmetic, Springer-G.T.M., 1973.

17a. C. Small, Waring's Problem, Math. Magazine, 50 (January 1977).

18. R. M. Stemmler, The ideal Waring theorem for exponents 401–200,000, Math. Comp., 18 (1964) 144–146.

19. H. E. Thomas, Jr., Waring's problem for twenty-two biquadrates, Trans. A.M.S. 193 (1974), 427–430.

20. L. Tornheim, Sums of n^{th} powers in fields of prime characteristic, Duke Math. J., 4 (1938) 359–362.

21. G. L. Watson, A proof of the seven cube theorem, J. London Math. Soc., 26 (1951) 153–156.

22. A. Weil, Number of solutions of equations in finite fields, Bull. A.M.S., 55 (1949) 497–508.

DEPARTMENT OF MATHEMATICS, QUEEN'S UNIVERSITY, KINGSTON, ONTARIO, CANADA.

SOME ERRORS IN APPLIED MATHEMATICS

WILLIAM C. WATERHOUSE

Mistakes in application of mathematics usually lie hidden in subject-area journals and escape the notice of the mathematical community. But the article *Signed digraphs and the energy crisis* **[3]** provides readers of the MONTHLY with two readily accessible examples. I want to analyze them here, not simply to criticize **[3]** but in hopes that the dissection will reveal something about the nature of applied mathematics. This paper can be understood without previous knowledge of **[3]**, and the specific examples discussed serve mainly as proof that the fallacies really do occur in practice.

1. The first error is a quite simple one. At the beginning of **[3]** is a signed digraph having, among other vertices, one labeled U (Energy Use) and one labeled R (Energy Price); there is a negatively-signed arrow from U to R. The authors say (p. 578): "According to the present system, this arc is $-$, because according to the present rate structure, the more you use, the less you pay (per kilowatt hour). It has been suggested that the rate structure should be inverted, and that large users should pay more rather than less. This strategy, known as inverting the rate structure, corresponds to changing the sign of arc (U, R) from $-$ to $+$."

Unfortunately, a little thought shows that this is nonsense, because the vertex U represents *total* energy use. (As confirmation that it does, note that the digraph has positively-signed arrows to U from "Population" and "Number of Factories.") Changing the sign of arc (U, R) has nothing to do with inverting the rate structure, because the distinction between large and small users is not in the picture.

2. The second error is less obvious and hence more interesting. The authors discuss signed digraphs, which indicate only whether each effect is augmenting or inhibiting; they point out (p. 579) that for some systems these may be the most detailed models available, particularly if some of the variables are hard to measure. They contrast signed digraphs with weighted digraphs, where the arrows are assigned numerical weights. For their detailed study, however, they include one theory in the other by treating signed digraphs as digraphs with weights ± 1 (cf. footnote, p. 589).

Unfortunately, this is fallacious; the results proved in that interpretation do not remain valid if one allows for the originally intended vagueness of the data in signed digraphs. To see this explicitly, even the simplest example will do: take two vertices joined to each other by positively-signed arrows. Theorems 5 and 6 of [3] assert, in the authors' interpretation, that a system with this signed digraph is pulse stable but value unstable. Yet if the plus signs would be weights $+(1/2)$ in a more detailed model, the same theorems show that the system is stable in both senses; and if they represent weights $+2$, it is unstable in both senses.

Indeed, it is not hard to see that the stability/instability results of [3] for signed digraphs are all factitious. Specifically, those who remember the definitions may enjoy proving the following statements: (I) For any signed digraph whatever, there is an assignment of weights with the given signs yielding a weighted digraph which is (in both senses) stable. (II) Consider a signed digraph which is stable in either of the senses of [3] and has at least one nonzero eigenvalue (nontrivial cycling effect). Then increasing the absolute value of the weights by an arbitrarily small amount will render it unstable.

3. Ecologists may wish to consider whether these errors affect our understanding of energy use in San Diego. For our purpose, however, it is more interesting simply to understand the nature of the errors themselves, since they are definitely mistakes in the application of mathematics rather than in mathematical technique. In a strictly formalized sense, the "mathematics" in [3] is correct; that is, the statements labeled "Theorem" are true, and the proofs presented for them are valid. Yet we have seen that the article does contain errors, errors lying not in the theorems but in the passage from theorem to practical interpretation. These, then, are specifically errors in *applied* mathematics.

But they are indeed errors in applied mathematics rather than in ecology or some other science. To see this we need only consider how they were shown to be erroneous. No crucial experiment contradicted a theory; no improved data on transportation were introduced; no deeper intuitive feeling for ecological systems suggested that something was wrong; in short, no empirical science at all was involved. The fallacies were isolated and refuted in exactly the same way that mistakes are caught in pure mathematics. Hence the mistakes are essentially mathematical themselves.

4. We thus see that the non-formalized parts of applied mathematics are subject to mathematical errors. But much more than this is true: the errors are of types familiar in mathematics, as we can quickly show with some historical examples. The first mistake, for instance, clearly comes from an imprecision in terms. Looking at a rate table, we cannot deny that for increased use the rate goes down; and one must bear a wary eye to see that this "increased use" is not quite the increase in use involved in the digraph. Quite similar unintentional ambiguity notoriously bedeviled the early study of analysis down to the time of Cauchy. Even he, who greatly clarified continuity, defined $f(x)$ to be a continuous function if the change in $f(x)$ approaches zero when the change in x does; the latent ambiguity of this eluded him, and he fell into fallacy through not distinguishing continuity from uniform continuity [1, p. 19 and pp. 124–5].

The second mistake has no simple name, but is equally familiar. It is easy, and for some purposes legitimate, to represent + and − by $+1$ and -1. But it is necessary to specify the limits within which such a representation is legitimate: stability properties of digraphs with weights ± 1 turn out not to yield corresponding statements for signed digraphs. The fallacy lies in tacitly supposing that a representation valid for a certain range of properties will continue to reflect properties of the original outside that range. And there is a famous example of exactly this mistake in the early study of the vibrating string: Euler at one time thought that a function like $x(\pi - x)$ could not be represented as a sine series on $[0, \pi]$, since the sum of the series would be an odd function and $x(\pi - x)$ was not [2, p. 236–7].

5. Most readers can probably match these historical examples with ones from their students' work, and indeed from their own research — we are hardly all better than Euler and Cauchy! Experience and training have taught us that only precise, persistent thinking can keep us from lapsing into error in mathematics. What the examples discussed here show is that exactly the same thinking is needed to

escape errors in application. It would be wrong, then, to suppose that applications are not mathematical when they are not fully formalized. The formalization now demanded in pure mathematics is of course invaluable as a guide for non-experts and a check against mistakes. But it simply displays the underlying correctness of thought; it does not create it. Few would maintain that all application of mathematics should be formalized. But it must be the object of sound mathematical thought.

6. Perhaps one practical conclusion might be drawn from this study of errors in application. Pure mathematicians sometimes look on those working in applications as a lesser breed without the Law, and students of applied mathematics sometimes have the impression that "rigor" has little to do with them. But it appears that in fact mathematicians of every kind need the same essential skills and training. Teachers of analysis know that students struggling to grasp the difference between continuity and uniform continuity, or figuring out just how a Fourier series works, are being forced to do some hard thinking. They are, obviously, learning information valuable and interesting in itself. But they are also familiarizing themselves with modes of thought which, as we have here seen in detail, may be exactly what they need in seemingly unrelated areas.

References

1. A. Cauchy, Leçons sur le Calcul Infinitésimal, 1823; Oeuvres II, 4 (Paris, 1899).

2. L. Euler, Remarques sur les mémoires précédens de M. Bernoulli, 1755; Opera Omnia II, 10, 233–254 (Bern, 1947).

3. F. S. Roberts and T. A. Brown, Signed digraphs and the energy crisis, this MONTHLY, 82 (1975) 577–593.

DEPARTMENT OF MATHEMATICS, PENNSYLVANIA STATE UNIVERSITY, UNIVERSITY PARK, PA 16802.

The authors of the original paper reply:

"We certainly agree with the main thesis, that many errors in applied mathematics are analogous to errors in pure mathematics. Also, the energy use vertex U must indeed be considered total energy use. This is easily rectified, and then the proper interpretation of the switch of sign on the arc (U, R) is that it is inverting the global rate structure, which at present increases energy price per kilowatt hour whenever total usage declines (as occurred during the energy crisis of 1973–74).

The other purported error rests on reading more into our paper than we say. It would certainly be wrong to make the assertion that, given a weighted digraph with unknown weights, we can determine its stability properties by studying the corresponding signed digraph. That we do not make this assertion should be evident from the definition of pulse process and from the theorems which we advance about stability, both of which refer explicitly to the weights. In addition, the simplicity of the example given by Professor Waterhouse should make it clear that we couldn't be making the assertion he says we are making. Indeed, as we ourselves have pointed out elsewhere (Ref. **10** of our paper), stability conclusions about weighted digraphs can even be wrong if one is not sure of the exact weights, for a small change in weights can change a stable weighted digraph into an unstable one, or vice versa. For this reason, we recommend a 'sensitivity analysis,' involving change of weights, be performed before the stability conclusions are accepted.

As to when stability and instability conclusions from a *signed* digraph analysis are meaningful, we note that they depend, among other things, on the assumption that the weights are all of unit magnitude. Even if this assumption is satisfied, a conclusion of stability (though not necessarily one of instability) will usually fail a sensitivity analysis, especially if all small changes in weight are permissible. If one has reason to question the assumption about unit weights — and one usually does — then one should try to obtain weights. Finally, as we pointed out on several occasions on pages 591 and 591–592, since the assumptions involved in pulse process modelling with signed (or even weighted) digraphs are very oversimplified, the conclusions should only be taken as suggestive and should be tested using analyses other than pulse processes."

MATHEMATICAL NOTES

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Laboratoire Calcul des Probabilités, Université de Paris, T.56, 4 Place Jussieu, 75–230 Paris, France.

AN APPLICATION OF HADAMARD'S INVERSE FUNCTION THEOREM TO ALGEBRA

WILLIAM B. GORDON

1. Introduction. In this note we shall show how the Inverse Function Theorem can be used to prove the well-known fact that euclidean N-space $\mathbf{E}^N$ cannot be endowed with the structure of a commutative division algebra when $N \geqq 3$. More precisely, we consider the possibility of defining an operation of "multiplication" $\mathbf{x}, \mathbf{y} \rightarrow \mathbf{xy}$ on $\mathbf{E}^N$ which for all scalars λ and vectors $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ satisfies the following four axioms:

(i) $$\mathbf{x}(\lambda \mathbf{y}) = (\lambda \mathbf{x})\mathbf{y} = \lambda \mathbf{xy},$$

(ii) $$\mathbf{x}(\mathbf{y} + \mathbf{z}) = \mathbf{xy} + \mathbf{xz},$$

(iii) $$\mathbf{xy} = 0 \text{ implies } \mathbf{x} = 0 \text{ or } \mathbf{y} = 0,$$

(iv) $$\mathbf{xy} = \mathbf{yx}.$$

We shall prove the following theorem:

THEOREM A. *For $N \geqq 3$, there is no operation of multiplication on* $\mathbf{E}^N$ *which satisfies* (i)-(iv).

A somewhat novel feature of our proof is that although we assume commutivity, we do not assume the associative law: $\mathbf{x}(\mathbf{yz}) = (\mathbf{xy})\mathbf{z}$. In fact, if the associative law were assumed in addition to (i)-(iv), then one could prove the existence of a unit $\mathbf{e}$ satisfying $\mathbf{xe} = \mathbf{ex} = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbf{E}^N$. (For let $\mathbf{a}$ be a fixed non-zero element of $\mathbf{E}^N$. Axiom (iii) implies that the linear map $\mathbf{x} \rightarrow \mathbf{ax}$ is non-singular, so that $\mathbf{ae} = \mathbf{a}$ for some vector $\mathbf{e}$. Hence, for any vector $\mathbf{b}$ we have $\mathbf{ab} = (\mathbf{ae})\mathbf{b} = \mathbf{a}(\mathbf{eb})$, from which it follows that $\mathbf{b} = \mathbf{eb}$.) Then $\mathbf{E}^N$ would be a finite field extension of $\mathbf{R}$ (identified with the set of all scalar multiples $\lambda \mathbf{e}$). From the Fundamental Theorem of Algebra it would then follow that $\mathbf{E}^N$ is $\mathbf{R}$ or $\mathbf{C}$.

For more general accounts of the problem concerning the existence of division algebras algebraic over the reals see e.g., [**1**], [**2**, p. 320], [**5**, p. 326–329], and [**6**].

2. The proof. Our proof will consist in showing that axioms (i)-(iv) imply that the map $\mathbf{x} \rightarrow \mathbf{x}^2$ is a homeomorphism when $N \geqq 3$, (which is obviously absurd since axiom (i) requires that $(-\mathbf{x})^2 = \mathbf{x}^2$). The proof of this assertion will depend in turn on the fact that the space $\mathbf{E}^N - \{p\}$ (euclidean space with a point removed) is simply connected when $N \geqq 3$. We shall use the following global version of the Inverse Function Theorem which is due to Hadamard. For details and references see [**3, 4**].

THEOREM B (Hadamard). *Let $f: M_1 \rightarrow M_2$ be a C^1 map between two connected N-dimensional manifolds whose Jacobian never vanishes, and which is "proper" in the sense that $f^{-1}(K)$ is compact whenever K is a compact subset of M_2. Suppose further that M_2 is simply-connected. Then f is a homeomorphism.*

To apply this theorem, let $M_1 = M_2 = \mathbf{E}^N - \{\mathbf{0}\}$. Axiom (iii) implies that the map $g(\mathbf{x}) = \mathbf{x}^2$ can be restricted to a map f from $\mathbf{E}^N - \{\mathbf{0}\}$ into itself, and one can easily show that f is continuous and proper. Using axiom (iv) to compute $df_{\mathbf{x}}(\mathbf{v})$ (the differential of f and $\mathbf{x}$ operating on $\mathbf{v}$) we get

$$df_{\mathbf{x}}(\mathbf{v}) = \lim_{h\to 0} \left\{\frac{1}{h}\left(f(\mathbf{x}+h\mathbf{v}) - f(\mathbf{x})\right)\right\} = \mathbf{xv} + \mathbf{vx} = 2\mathbf{xv}.$$

Hence axiom (iii) implies that $df_{\mathbf{x}}$ is non-singular for all $\mathbf{x}$ in $\mathbf{E}^N - \{\mathbf{0}\}$, so that f is a homeomorphism.

References

1. C. W. Curtis, The four and eight square problem and division algebras, in Studies in Modern Algebra, MAA Studies in Math., Vol 2, (A. A. Albert, ed.), Prentice Hall, Englewood Cliffs, N. J., 1968.

2. S. Eilenberg and N. Steenrod, Foundations of Algebraic Topology, Princeton U. Press, Princeton, N.J., 1952.

3. W. B. Gordon, On the diffeomorphisms of euclidean space, this MONTHLY, 79 (1972) 755–759.

4. ———, Addendum to the above, this MONTHLY, 80 (1973) 674.

5. I. N. Herstein, Topics in Algebra, Blaisdell, Waltham, Mass., 1964.

6. R. S. Palais, The classification of real division algebras, this MONTHLY, 75 (1968) 366–368.

RADAR DIVISION, NAVAL RESEARCH LABORATORY, CODE 5308, WASHINGTON, DC 20375.

SOME THEOREMS ON POLYNOMIALS OVER A FINITE FIELD

L. CARLITZ

1. Let $F = GF(q)$ denote the finite field of order q. For polynomials $f(x) \in F[x]$ in a single indeterminate it is easy to prove [**1**] that $f(x+a) = f(x)$ for all $a \in F$ if and only if

$$f(x) = \sum_k c_k (x^q - x)^k, \tag{1}$$

where the c_k are arbitrary elements of F. Indeed, for arbitrary $f(x) \in F[x]$, we have

$$f(x) = \sum_k c_k(x)(x^q - x)^k,$$

where $c_k(x) \in F[x]$, $\deg c_k(x) < q$ $(k = 0, 1, 2, \ldots)$. It follows since the $c_k(x)$ are uniquely determined by $f(x)$, that from $f(x+a) = f(a)$,

$$c_k(x+a) = c_k(x) \qquad (a \in F)$$

and this in turn implies that $c_k(x)$ is free of x. Conversely it is clear that for $f(x)$ given by (1) we have $f(x+a) = f(x)$.

In the case of two indeterminates let $f(x, y) \in F[x, y]$ and assume that

$$f(x+a, y+b) = f(x, y) \tag{2}$$

for all a, $b \in F$. Then it is easy to prove, as in the case of a single indeterminate, that (2) holds if and only if

$$f(x, y) = \sum_{j,k} c_{j,k}(x^q - x)^j (y^q - y)^k, \tag{3}$$

where the $c_{j,k}$ are independent of x, y.

A more interesting question, however, is the following. What polynomials $f(x, y) \in F[x, y]$ satisfy

$$f(x+a, y+a) = f(x, y) \tag{4}$$

for all $a \in F$?

Let $f(x, y)$ be a polynomial that satisfies (4) and put

$$f(x, y) = \sum_{j,k} c_{j,k}(x, y)(x^q - x)^j (y^q - y)^k, \tag{5}$$

where the $c_{j,k}(x, q)$ are polynomials of degree $\leqq q-1$ in each indeterminate. Since the $c_{j,k}(x, q)$ are uniquely determined by (5), it follows from (4) that

$$c_{j,k}(x + a, y + a) = c_{j,k}(x, y) \qquad (a \in F). \tag{6}$$

We now prove the following:

LEMMA. *Let $f(x, y)$ be a polynomial $\in F[x, y]$ of degree $\leqq q-1$ in each indeterminate. Then $f(x, y)$ satisfies*

$$f(x + a, y + a) = f(x, y) \qquad (a \in F) \tag{7}$$

if and only if

$$f(x, y) = \phi(x - y), \tag{8}$$

for some $\phi(x) \in F[x]$, deg $\phi(x) < q$.

Proof. Put $\psi(x, y) = f(x - y, 0)$. Then

$$\psi(a, b) = f(a - b, 0) = f(a, b), \tag{9}$$

by (7). Since $\phi(x, y)$ and $f(x, y)$ are of degree $< q$ in each indeterminate, it follows from (9) that

$$f(x, y) = \psi(x, y) = f(x - y, 0). \tag{10}$$

Conversely, it is clear that (8) implies (7).

Thus we have proved

THEOREM 1. *Let $f(x, y) \in F[x, y]$. Then $f(x + a, y + a) = f(x, y)$ for all $a \in F$ if and only if*

$$f(x, y) = \sum_{j,k} c_{j,k}(x - y)(x^q - x)^j (y^q - y)^k,$$

where the $c_{j,k}(x)$ are arbitrary polynomials $\in F[x]$ of degree $< q$.

The extension to more than two indeterminates is almost immediate. For example, for three indeterminates, we have the following

THEOREM 2. *Let $f(x, y, z) \in F[x, y, z]$. Then $f(x + a, y + a, z + a) = f(x, y, z)$ for all $a \in F$ if and only if*

$$f(x, y, z) = \sum_{i,j,k} c_{i,j,k}(x - z, y - z)(x^q - x)^i (y^q - y)^j (z^q - z)^k,$$

where the $c_{i,j,k}(x, y)$ are arbitrary polynomials $\in F[x, y]$ of degree $< q$ in each indeterminate.

Theorem 1 may be extended in another direction: What polynomials $f(x, y, z) \in F[x, y, z]$ satisfy

$$f(x + a, y + b, z + c) = f(x, y, z) \tag{11}$$

for all $a, b, c \in F$ such that

$$a + b + c = 0? \tag{12}$$

To answer this question put

$$f(x, y, z) = \sum_{i,j,k} c_{i,j,k}(x, y, z)(x^q - x)^i (y^q - y)^j (z^q - z)^k, \tag{13}$$

where the $c_{i,j,k}(x,y,z)$ are uniquely determined polynomials $\in F[x,y,z]$ of degree $< q$ in each indeterminate. It then follows from (11) and (13) that

$$c_{i,j,k}(x+a,y+b,z+c)=c_{i,j,k}(x,y,z) \tag{14}$$

for all $a,b,c \in F$ such that $a+b+c=0$. Thus we have reduced the problem to the case of $f(x,y,z)$ with restricted degrees.

By (11) we have

$$f(x+a,y-a,z)=f(x,y,z) \qquad (a \in F). \tag{15}$$

By a slight extension of Theorem 1, (15) holds for all $a \in F$ if and only if $f(x,y,z)=f_1(x+y,z)$, for some $f_1(x,z)\in F[x,z]$ of degree $< q$ in each indeterminate. Then (11) gives

$$f_1(x+c,z-c)=f_2(x+z),$$

for some $f_2(x)\in F[x]$ of degree $< q$. Thus if $f(x,y,z)$ is of degree $< q$ in each indeterminate, then it satisfies (11) if and only if $f(x,y,z)=\phi(x+y+z)$, where $\phi(x)\in F[x]$, $\deg \phi(x)<q$.

We may now state

THEOREM 3. *Let $f(x,y,z)\in F[x,y,z]$. Then $f(x,y,z)$ satisfies*

$$f(x+a,y+b,z+c)=f(x,y,z)$$

for all $a,b,c\in F$ such that $a+b+c=0$ if and only if

$$f(x,y,z)=\sum_{i,j,k} c_{i,j,k}(x+y+z)(x^q-x)^i(y^q-y)^j(z^q-z)^k,$$

where the $c_{i,j,k}(x)$ are arbitrary polynomials $\in F[x]$ of degree $< q$.

The extension to an arbitrary number of indeterminates is immediate.

2. The theorems proved above can be extended further in a different direction. Let F_1 denote the field $F(t)$ consisting of all rational functions in the indeterminate t over the field F. By Theorem 2.1 of [**1**] if $f(x)\in F_1[x]$, then

$$f(x+A(t))=f(x) \qquad (\deg A(t)<m) \tag{16}$$

for all polynomials $A(t)\in F_1$ of degree $< m$ if and only if

$$f(x)=\sum_k c_k(\psi_m(x))^k, \tag{17}$$

where the $c_k\in F_1$ and are independent of x and $\psi_m(x)=\Pi_{\deg A(t)<m}(x-A(t))$.

The polynomial $\psi_m(x)=\Sigma_{k=0}^m(-1)^{m-k}\begin{bmatrix} m\\ k\end{bmatrix}x^{q^k}$, where

$$\begin{bmatrix} m\\ k\end{bmatrix}=\frac{F_m}{F_kL_{m-k}^{qk}},\quad F_m=\prod_{j=0}^{m}(t^{q^m}-t^{q^j}),\quad L_m=\prod_{j=1}^{m}(t^{q^j}-t);$$

moreover, F_m is the product of the monic polynomials $A(t)$ of degree m and L_m is the least common multiple of the polynomials of degree m.

Thus, for $f(x,y)\in F_1[x,y]$, we may consider

$$f(x+A(t),y+A(t))=f(x,y) \tag{18}$$

for all $A(t)$ of degree $< m$. For $m=1$, (18) reduces to (4). The proof of Theorem 1 is easily carried over to this more general situation. We state

THEOREM 4. *Let* $f(x, y) \in F_1[x, y]$, *where* $F_1 = F(t)$. *Then* $f(x, y)$ *satisfies* (18) *for all polynomials* $A(t) \in F_1$ *of degree* $< m$ *if and only if*

$$f(x, y) = \sum_{j,k} c_{j,k}(x - y)(\psi_m(x))^j(\psi_m(y))^k,$$

where the $c_{j,k}(x)$ *are arbitrary polynomials* $\in F_1[x]$ *of degree* $< q^m$.

Since $\psi_1(x) = x^q - x$, it is evident that, for $m = 1$, Theorem 4 reduces to Theorem 1.
It is unnecessary to state the generalized version of Theorem 2. As for Theorem 3 we have

THEOREM 5. *Let* $f(x, y, z) \in F_1[x, y, z]$. *Then* $f(x, y, z)$ *satisfies*

$$f(x + A(t), y + B(t), z + C(t)) = f(x, y, z)$$

for all polynomials $A(t)$, $B(t)$, $C(t) \in F_1$ *of degree* $< q$ *and such that* $A(t) + B(t) + C(t) = 0$, *if and only if*

$$f(x, y, z) = \sum_{i,j,k} c_{i,j,k}(x + y + z)(\psi_m(x))^i(\psi_m(y))^j(\psi_m(z))^k,$$

where the the $c_{i,j,k}(x)$ *are arbitrary polynomials* $\in F_1[x]$ *of degree* $< q^m$.

REMARK. It may be worth while remarking that there are numerous applications of finite fields to cryptanalysis, coding theory, statistics and electrical engineering.

Supported in part by NSF grant GP-37924X.

Reference

1. L. Carlitz, Finite sums and interpolation formulas over GF $[p^n, x]$, Duke Math. J., 15 (1948) 1001–1012.

DEPARTMENT OF MATHEMATICS, DUKE UNIVERSITY, DURHAM, NC 27706.

SOME REMARKS ON THE IRRATIONAL AND RATIONAL NUMBERS

CARL EBERHART

The first-year graduate student usually learns in his topology course that the Cantor middle third set is characterized topologically by the properties *metric*, *compact*, *totally disconnected*, and *has no isolated point* (which is the same as: every finite subset has void interior). Similar characterizations of the irrational numbers and the rational numbers (with the topology inherited from the line) do not seem to be as well-known although they are in the literature. F. Hausdorff in [**2**, p. 157] characterized the irrationals by the properties *separable metric*, *topologically complete* (i.e., admits a complete metric), *zero-dimensional*, and *nowhere locally compact* (i.e., every compact subset has void interior). The rationals are characterized by W. Sierpinski in [**5**] by the properties *metric*, *countable*, and *has no isolated point.*

In this note we will prove a theorem from which Hausdorff's characterization and Sierpinski's characterization are obtained as corollaries. Also we give two applications of these characterizations.

THEOREM 1. *Suppose X and Y are separable metric spaces which are topologically complete, zero-dimensional, and nowhere locally compact. Let D and E be countable dense subsets of X and Y respectively. Then there is a homeomorphism h from X onto Y taking D onto E.*

Proof. The homeomorphism h will be "induced" by a 1-1 function H from a suitably chosen base

$\mathcal{B}$ for the topology on X onto a suitably chosen base $\mathcal{C}$ for the topology on Y. (This technique is used in [3; pp. 97–100] to characterize the Cantor set.)

Give X a complete metric and construct a sequence $\mathcal{B}_1, \mathcal{B}_2, \ldots$ of covers of X such that for $n = 1, 2, \ldots,$

(i) the members of $\mathcal{B}_n$ are pairwise disjoint open and closed sets of diameter $< 1/n$, and

(ii) $\mathcal{B}_1$ is infinite, $\mathcal{B}_{n+1}$ refines $\mathcal{B}_n$, and each member of $\mathcal{B}_n$ contains infinitely many members of $\mathcal{B}_{n+1}$.

We may construct the covers $\mathcal{B}_1, \mathcal{B}_2, \ldots$ as follows: Since X is not compact the complete metric on X is not totally bounded so choose $0 < \varepsilon < 1$ so that no finite collection of sets of diameter $< \varepsilon$ covers X. Now choose a countable collection $U_1, U_2, \ldots$ of open and closed sets of diameter $< \varepsilon$ covering X and let $\mathcal{B}_1$ consist of all the sets $U_{n+1} \setminus (U_1 \cup \cdots \cup U_n)$ which are nonvoid. To construct $\mathcal{B}_2$, note that each member B of $\mathcal{B}_1$ is complete but not compact with respect to the metric on X and so it too can be written as the union of an infinite collection of pairwise disjoint open and closed sets of diameter $< \frac{1}{2}$. Form such a collection for each member of $\mathcal{B}_1$ and let $\mathcal{B}_2$ be the union of all these collections. The rest of the $\mathcal{B}_n$'s are constructed in the same manner. The set $\mathcal{B} = \bigcup_{n=1}^{\infty} \mathcal{B}_n$ is the desired base for the topology on X.

Notice that for each n, each point $x \in X$ lies in a unique member of $\mathcal{B}_n$. Denote this by $B_n(x)$. Thus we have

(iii) $B_1(x) \supset B_2(x) \supset \cdots \supset \bigcap_{n=1}^{\infty} B_n(x) = \{x\}$,

(iv) each tower $B_1 \supset B_2 \supset \cdots$ where $B_n \in \mathcal{B}_n$ determines a unique point of X.

Now fix an enumeration of D and notice that each $B \in \mathcal{B}$ contains a first point of D, that is, a point of D with smallest subscript. Denote this point by $D(B)$ and observe that if $B \in \mathcal{B}_n$, then for $m \geqq n$ we have

$$\text{(v)} \qquad D(B_m(D(B))) = D(B);$$

that is, $D(B)$ is also the first point of D in each member of $\mathcal{B}$ containing it which is contained in B.

Construct a base $\mathcal{C} = \bigcup_{n=1}^{\infty} \mathcal{C}_n$ for the topology on Y and enumerate E exactly as we did for X so that properties analogous to (i)–(v) are satisfied.

The function H is defined step by step as follows: Since $\mathcal{B}_1$ and $\mathcal{C}_1$ are (countably) infinite, we let H_1 be any 1-1 function from $\mathcal{B}_1$ onto $\mathcal{C}_1$. Assuming a 1-1 function H_n from $\mathcal{B}_n$ onto $\mathcal{C}_n$ has been given let H_{n+1} be any 1-1 function from $\mathcal{B}_{n+1}$ onto $\mathcal{C}_{n+1}$ which is compatible with H_n in the sense that for each $B \in \mathcal{B}_n$,

(a_n) H_{n+1} takes the members of $\mathcal{B}_{n+1}$ contained in B to the members of $\mathcal{C}_{n+1}$ contained in $H_n(B)$, and

(b_n) H_{n+1} takes the member of $\mathcal{B}_{n+1}$ containing $D(B)$ to the member of $\mathcal{C}_{n+1}$ containing $E(H_n(B))$; i.e.,

$$H_{n+1}(B_{n+1}(D(B))) = C_{n+1}(E(H_n(B))).$$

The function $H: \mathcal{B} \to \mathcal{C}$ given by $H(B) = H_n(B)$ if $B \in \mathcal{B}_n$ then satisfies (using v),

(a) if $B, B' \in \mathcal{B}$, then $H(B) \subset H(B')$ if and only if $B \subset B'$,

(b) if $B \in \mathcal{B}_N$, then for $n \geqq N$, $H(B_n(D(B))) = C_n(E(H(B)))$.

Properties (iii), (iv) and (a) guarantee that the equation

$$\{h(x)\} = \bigcap_{n=1}^{\infty} H(B_n(x))$$

defines a 1-1 function h from X onto Y. The continuity of h and h^{-1} follows from properties (i), (ii) and (iii). The details are similar to those in [3] and are omitted. To see that h takes D onto E, let $d \in D$ and choose N so large that $D(B_N(d)) = d$. Then

$$h(d) = \bigcap_{n=N}^{\infty} H(B_n(d)) = \bigcap_{n=N}^{\infty} H(B_n(D(B_N(d)))) = \bigcap_{n=N}^{\infty} C_n(E(H(B_N(d)))) = E(H(B_N(d))) \in E$$

and thus $h(D) \subset E$, Now let $e \in E$ and choose N so large that $E(C_N(e)) = e$. Let $B = H^{-1}(C_N(e))$ and note that

$$h(D(B)) = \bigcap_{n=N}^{\infty} H(B_n(D(B))) = \bigcap_{n=N}^{\infty} C_n(E(H(B))) = \bigcap_{n=N}^{\infty} C_n(E(C_N(e))) = \bigcap_{n=N}^{\infty} C_n(e) = e$$

and so $h(D) \supset E$. This completes the proof.

The space of irrational numbers is easily seen to be separable metric, zero-dimensional, and nowhere locally compact. The fact that it admits a complete metric follows from the fact that it is a G_δ set in a complete metric space [**4**; p. 408]. Thus we have Hausdorff's characterization of the irrationals as a corollary to Theorem 1. Actually Theorem 1 also shows that the irrationals are *countable dense homogeneous*; that is, any two of its countable dense subsets are homeomorphic by means of a self homeomorphism of the irrationals. The notion of countable dense homogeneous is due to Ralph Bennett [**1**], and provides the means of obtaining Sierpinski's characterization of the rationals from Theorem 1.

COROLLARY 2. *Every countable metric space with no isolated point is homeomorphic with the rationals.*

Proof. Let X be a countable metric space. Then X is zero dimensional and hence can be regarded as a subspace of the Cantor set C [**4**; p. 285]. Since X has no isolated point, the closure of X in C is compact, totally disconnected and has no isolated point, hence is homeomorphic with C. So we can assume that X is a dense subset of C. Let E be the set of "endpoints" of C, that is, $E = \{0, 1, 1/3, 2/3, \ldots\}$, and let A be any countable dense subset of C such that $A \cap (X \cup E) = \emptyset$. Note that $C \backslash A$ is separable metric and zero dimensional. It is topologically complete since it is a G_δ-set in C. That it is nowhere locally compact follows from the more general fact that if A is a dense subset of a Hausdorff space X such that $X \backslash A$ is dense, then A is nowhere locally compact. Thus $C \backslash A$ is homeomorphic with the irrationals, and since E and X are countable dense subsets of $C \backslash A, X$ is homeomorphic with E. Thus all countable metric spaces with no isolated point are homeomorphic.

It is an interesting consequence of the above proof that if D is any countable dense subset of the Cantor set, then D is homeomorphic with the rational numbers and $C \backslash D$ is homeomorphic with the irrational numbers. This fact is generalized in the next theorem (compare with [**4**, pp. 437–442]).

THEOREM 3. *Let X be a separable metric space which is topologically complete and has no isolated point. Let D be a countable dense subset of X. Then D is homeomorphic with the rationals and $X \backslash D$ contains a dense subset homeomorphic with the irrationals.*

Proof. D has no isolated point since X does not. Hence D is homeomorphic with the rationals (the completeness of X is not needed here). To construct the dense copy Q of the irrationals in $X \backslash D$, let $\mathscr{B} = \{U_1, U_2, \ldots\}$ be a countable base for X and let B_n be the boundary of U_n for each n. Enumerate $D = \{d_1, d_2, \ldots\}$ and note that the sets $V_n = X \backslash (B_n \cup \{d_1, \ldots, d_n\})$ are dense open subsets of X. Hence by the Baire category theorem, the set $Q = \bigcap_{n=1}^{\infty} V_n$ is dense in X. It is also topologically complete, zero-dimensional, nowhere locally compact, and is contained in $X \backslash D$.

The term *rational curve* is applied to those compact connected metric spaces which have a base consisting of open sets with countable boundaries [**6**]. By a *compactification* of a metric space X by a space Y we mean a compact metric space Z which contains a dense homeomorphic image of X whose complement is homeomorphic with Y. Rational curves are characterized by

THEOREM 4. *Let X be a compact connected metric space. Then X is a rational curve if and only if X is a compactification of the irrationals by the rationals.*

Proof. If X is a rational curve, then in the proof of Theorem 3 choose the base $\mathscr{B}$ so that B_n is countable for all n and note that $(\bigcup_{n=1}^{\infty} B_n) \cup D$ is homeomorphic with the rationals. Conversely suppose X is the disjoint union of the rationals R with the irrationals I, and let $x \in U$, where U is open in X. Note that $I \cup \{x\}$ is zero dimensional. Hence there is an open set V in X such that $x \in V \subset U$ and the boundary of $V \cap (I \cup \{x\})$ in $I \cup \{x\}$ is empty. So the boundary of V is contained in R, hence is countable. Thus X is a rational curve.

References

1. Ralph Bennett, Countable dense homogeneous spaces, Fund. Math., 74 (1972) 189–194.

2. F. Hausdorff, Die schlichten stetigen Bilder des Nullraums, Fund. Math., 29 (1937) 151–158.

3. J. Hocking and G. Young, Topology, Addison Wesley, Reading, Mass., 1961.

4. K. Kuratowski, Topology, vol. 1, New York and Warszawa, 1966.

5. Waclaw Sierpinski, Sur une propriété topologique des ensembles dénombrables dense en soi, Fund. Math., 1 (1920) 11–16.

6. G. T. Whyburn, Analytic Topology, A.M.S. Colloq. Publ. No. 28, 1942.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF KENTUCKY, LEXINGTON, KY 40506.

DIVISIBILITY OF BINOMIAL COEFFICIENTS BY THEIR ROW NUMBER

HEIKO HARBORTH

Dedicated to Professor Richard K. Guy on his sixtieth Birthday

In this note we consider the binomial coefficients $\binom{m}{k}$ and ask how many of them are divisible by m. This question may arise in connection with some necessary and sufficient conditions for primality (see [**2**] and [**5**]).

In Pascal's triangle we find $\binom{m}{k}$ in the kth position of the mth row, $0 \leqq k \leqq m$. If we write 0 in case $\binom{m}{k} \equiv 0 \pmod{m}$, and $*$ otherwise, we get a certain pattern. The first part of this pattern is shown in Figure 1.

```
                                        *
                                      0   0
                                    *   0   *
                                  *   0   0   *
                                *   0   *   0   *
                              *   0   0   0   0   *
                            *   0   *   *   *   0   *
                          *   0   0   0   0   0   0   *
                        *   0   *   0   *   0   *   0   *
                      *   0   0   *   0   0   *   0   0   *
                    *   0   *   0   0   *   0   0   *   0   *
                  *   0   0   0   0   0   0   0   0   0   0   *
                *   0   *   *   *   0   0   0   *   *   *   0   *
              *   0   0   0   0   0   0   0   0   0   0   0   0   *
            *   0   *   0   *   0   *   *   *   0   *   0   *   0   *
          *   0   0   *   0   *   *   0   0   *   *   0   *   0   0   *
        *   0   *   0   *   0   *   0   *   0   *   0   *   0   *   0   *
      *   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   *
    *   0   *   *   0   0   *   0   0   *   0   0   *   0   0   *   *   0   *
```

FIG. 1

Let $A(N)$ denote the number of symbols 0 within the first N symbols, counted row by row from left to right. We then consider the sequence $\{A(N)/N\}$ and prove

$$\lim_{N\to\infty}\frac{A(N)}{N}=1, \tag{1}$$

that is in other words:

THEOREM. *Almost all binomial coefficients* $\binom{m}{k}$ *are divisible by their row number* m.

For the proof we need three known results. It was shown in **[3]** and **[4]**, that almost all binomial coefficients are divisible by any fixed number t. That means

$$Z(r)=o\left(\binom{r+1}{2}\right)=o(r^2), \tag{2}$$

where $Z(r)$ denotes the number of coefficients $\binom{m}{k}$ not being divisible by t, and with $m\leqq r-1$. In **[1]**, p. 245 we find

$$\sum_{i=1}^{\infty}\frac{1}{i^2}=\frac{\pi^2}{6}. \tag{3}$$

Further we use

$$\sum_{i=1}^{r}\phi(i)=\frac{3r^2}{\pi^2}+O(r\log r)=\frac{3r^2}{\pi^2}+o(r^2) \tag{4}$$

from **[1]**, p. 268 ($\phi(i)$ denotes Euler's totient function).

Now for $(m,k)=t$ we have

$$\binom{m}{k}=\frac{m}{k}\binom{m-1}{k-1}\equiv 0\ (\mathrm{mod}\ m),\quad\text{if}\quad \binom{m-1}{k-1}\equiv 0\ (\mathrm{mod}\ t). \tag{5}$$

With $m=vt$, $k=wt$, $1\leqq w\leqq v$, the relation $(m,k)=t$ is seen to be equivalent to $(v,w)=1$. Thus for $m=vt\leqq n$ there are

$$F_n(t)=\sum_{v=1}^{[n/t]}\phi(v) \tag{6}$$

binomial coefficients $\binom{m}{k}$ with the property $(m,k)=t$ (excluding only $\binom{1}{0}$ for $t=1$).

Then

$$\binom{m-1}{k-1}=\binom{vt-1}{wt-1}=\binom{v-1}{w-1}\prod_{i=0}^{w-1}\prod_{j=1}^{t-1}\frac{(v-i)t-j}{it-j}. \tag{7}$$

Since all numbers in the numerator and denominator of (7), after being divided by (t,j), are prime to the modulus t, we obtain

$$\binom{m-1}{k-1}\equiv 0\ (\mathrm{mod}\ t),\quad\text{if and only if}\quad \binom{v-1}{w-1}\equiv 0\ (\mathrm{mod}\ t). \tag{8}$$

Thus for $m=vt\leqq n$, that is $v-1\leqq[n/t]-1$, there are $Z([n/t])$ coefficients

$$\binom{m-1}{k-1}=\binom{vt-1}{wt-1},$$

which are not divisible by t.

Using (6), (4), (3), and (2) we now conclude for all N with $\binom{n+2}{2} \leqq N < \binom{n+3}{2}$

$$\begin{aligned} \frac{A(N)}{N} &\geqq A\left(\binom{n+2}{2}\right) \Big/ \binom{n+3}{2} \geqq \frac{2}{(n+2)(n+3)} \sum_{t=1}^{n} \left\{ F_n(t) - Z\left(\left[\frac{n}{t}\right]\right) \right\} \\ &= \frac{2}{(n+2)(n+3)} \sum_{t=1}^{n} \left\{ \frac{3}{\pi^2} \left[\frac{n}{t}\right]^2 + o\left(\frac{n^2}{t^2}\right) \right\} \\ &= \frac{2n^2}{(n+2)(n+3)} \sum_{t=1}^{n} \left\{ \frac{3}{\pi^2 t^2} + o\left(\frac{1}{t^2}\right) \right\} \\ &= \frac{2n^2}{(n+2)(n+3)} \left\{ \frac{3}{\pi^2} + o(1) \right\} \sum_{t=1}^{n} \frac{1}{t^2} \xrightarrow[(n \to \infty)]{} 1. \end{aligned} \tag{9}$$

By (9) we have got (1), and so the Theorem is proved.

References

1. G. H. Hardy and E. M. Wright, Introduction to the Theory of Numbers, Oxford University Press (fourth ed.), London, 1960.

2. H. B. Mann and D. Shanks, A necessary and sufficient condition for primality, and its source, J. Combinatorial Theory (A), 13 (1972) 131–134.

3. D. Singmaster, Notes on binomial coefficients III — Any integer divides almost all binomial coefficients, J. London Math. Soc. (2), 8 (1974) 555–560.

4. H. Harborth, Über die Teilbarkeit im Pascal-Dreieck, Math.-Phys. Semesterberichte, 22 (1975) 13–21.

5. ———, Ein Primzahlkriterium nach Mann und Shanks, Arch. Math. (Basel), 27 (1976) 290–294.

TECHNISCHE UNIVERSITÄT, D 33 BRAUNSCHWEIG, WEST GERMANY.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics and Statistics, The University of Calgary, Calgary, Alberta, Canada, T2N 1N4. (From July 1976 to June 1977: Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, England.)

CONTINUED FRACTIONS IN ALGEBRAIC NUMBER FIELDS

DAVID ROSEN

Is it possible to devise a continued fraction that represents uniquely all real numbers, so that the finite continued fractions represent the elements of an algebraic number field, and conversely, every element of the number field is represented by a finite continued fraction? We shall call this the **representation property**. The continued fraction representation is useful if, in addition, the elements of the number field can be used to approximate "irrational" numbers (i.e., non-field elements). We shall say that the number field has the **approximation property** if for every "irrational" α

$$\left| \alpha - \frac{P}{Q} \right| < \frac{1}{kQ^2}$$

is satisfied by infinitely many rational elements P/Q of the number field and k is a positive fixed constant. The model we wish to emulate is the rational field represented by the regular simple continued fractions.

The object of this note is to accomplish two things. We first wish to give one example of an algebraic number field which has both the representation and approximation properties. We then present a class of number fields which have the approximation property and *almost* have the representation property. The *almost* leads to the research problem and conjectures.

The example: The algebraic number field with the desired properties is the familiar field generated by the golden number $\tau = 2\cos\pi/5 = (1+\sqrt{5})/2$. The elements of the number field have the form

$$\frac{a+b\tau}{c+d\tau},\quad a, b, c, d \text{ integers, } c, d \text{ not both } 0. \tag{1}$$

The continued fractions which will be called **τ-fractions** have the form

$$r_0 + \frac{\varepsilon_1}{r_1\tau +}\,\frac{\varepsilon_2}{r_2\tau +}\cdots,$$

where $\varepsilon_i = \pm 1$, r_0 is any integer and the other r_i are positive integers.

The representation is unique if the following rule is observed. If $r_i\tau + \varepsilon_{i+1} < 1$, then $r_{i+1} \geqq 2$. For example, $\tau - (1/2\tau)$ is permissible, but not $\tau - 1/\tau$, unless it stands by itself as a representation of 1. Since τ is a root of the polynomial

$$\tau^2 - \tau - 1 = 0 \tag{2}$$

the convergents of the continued fraction, when evaluated, give rise to the quotients of polynomials which can be reduced using (2) to the form in (1). Thus

$$2\tau - \frac{1}{2\tau} = \frac{4\tau^2-1}{2\tau} = \frac{4\tau+3}{2\tau},$$

$$\tau + \frac{-1}{2\tau +}\,\frac{-1}{\tau} = \tau + \frac{-\tau}{2\tau+1} = \frac{2\tau+2}{2\tau+1} = \frac{2(\tau+1)}{2\tau+1} = \frac{2}{\tau},$$

since $\tau + 1 = \tau^2$, and $2\tau + 1 = \tau(\tau+1)$.

In **[2]** it is shown that the τ-fractions and the **λ_q-fractions** arising from the number fields generated by $\lambda_q = 2\cos\pi/q$, q an odd positive number $\geqq 3$, have the approximation property. It is also shown that every real number has a unique representation if the following conditions are imposed:

1. Let $h = [(q-1)/2]$ ($[x]$ is the greatest integer $\leqq x$); then $r_i\lambda + \varepsilon_{i+1} < 1$ is satisfied for no more than h consecutive values of i.

2. If q is odd and $r_i\lambda + \varepsilon_{i+1} < 1$ is satisfied for h consecutive values of i, then $r_{i+h} \geqq 2$.

These are the main conditions, but two more technical conditions are needed to handle the tail of the finite fractions. These are given on p. 555 of **[2]**. The λ_q-fraction is then called a **reduced** λ_q-fraction.

The only fact not proved in **[2]** is that every member of the number field has a finite λ_q-fraction representation. In **[1]** Leutbecher supplies a proof for the τ-fraction; he shows that every element of the field $R(\tau)$ has a finite τ-fraction representation. The results of **[2]** and **[1]** therefore together supply one example.

What can one say about the other number fields $2\cos\pi/q$, q odd? A generalization of Leutbecher's theorem, combined with the results of **[2]**, is all that is needed to make these fields into examples. Unfortunately, Leutbecher's result doesn't seem to generalize.

The λ_q-fraction came about in trying to answer questions concerning the arithmetic character of the coefficients of linear fractional transformations associated with certain properly discontinuous groups called Hecke groups. Each λ_q field generates a group of transformations $V(z) =$

$(\alpha z+\beta)/(\gamma z+\delta)$, $\alpha,\beta,\gamma,\delta\in\lambda_q$ number field, which take the upper half of the complex plane into itself. For even q, the λ_q number fields cannot qualify as examples because, as it is shown in [2], the algebraic number 1 has an infinite λ_q-fraction representation. If for any other odd q, the λ_q number field qualifies as an example, it would be true that, for that field, the parabolic or cusp points (i.e., the points $-\delta/\gamma$ on the real line that are congruent to ∞ by $V(z)$) are precisely the elements of the field — an important result for the Hecke groups. For $q=3$, the associated Hecke group, denoted by Γ_3 is the familiar modular group. The Hecke group Γ_5 is associated with the τ field of the above example. The case $q=5$ may be special because τ generates a quadratic field. A tantalizing conjecture is that each of the λ_q number fields, q odd and $\geqq 7$ is an example, for this would imply that the parabolic points for the associated Hecke groups are the elements of the field.

Except for the papers mentioned below, no other texts or papers appear to consider the λ_q fraction or treat this problem.

References

1. Armin Leutbecher, Über die Heckeschen Gruppen G(λ), Abh. Math. Sem., Univ. Hamburg, 31 (1967) 199–205; MR 37 #4018.

2. David Rosen, A class of continued fractions associated with certain properly discontinuous groups, Duke Math. J., 21 (1954) 549–564; MR 16–458.

DEPARTMENT OF MATHEMATICS, SWARTHMORE COLLEGE, SWARTHMORE, PA 19081.

NONCANONICAL FACTORIZATION OF A PERMUTATION

J. L. BRENNER AND J. RIDDELL

It is well known that the alternating group $\mathscr{A}_n$ is generated by the 3-cycles. It is less well recognized that this is an inefficient way to generate $\mathscr{A}_n$, in the sense that no fewer than $[n/2]$ 3-cycles must be multiplied to obtain the permutation $(1\,2)\,(3\ldots n)$ (if n is even) or the permutation $(1\,2\ldots n)$ (if n is odd).

PROBLEM. *What permutations P in $\mathscr{A}_n$ can be expressed as the product of a permutation of period k by a permutation of period l?*

Remarks on this problem follow.

(i) If $P=P_1P_2$ and P_1,P_2 both have period 2, then P is obviously conjugate to its inverse. (It is known that if $n\neq 1,2,5,6,10,14$, $\mathscr{A}_n$ has a permutation that is not conjugate to its inverse.) The converse assertion is:

If $P\in\mathscr{A}_n$, and P is conjugate to its inverse, then P is expressible as the product of two permutations (in $\mathscr{A}_n$) of period 2. (Although this is not immediate, the proof is reasonably elementary. See [2].)

(ii) If $n>n_0=n_0(k)$, $k>2$, every permutation $P\in\mathscr{A}_n$, $P\neq 1$, can be written as the product of a permutation of period k by a permutation of period 2 (in $\mathscr{A}_n$).

(iii) If $n\geqq 3$, every permutation $P\in\mathscr{A}_n$ can be written as the product of two factors each of period 3.

(iv) If $n>11$, every permutation $P\in\mathscr{A}_n$ is expressible as the product of two permutations of period 5. For example, in $\mathscr{A}_{12}$,

$$(0\;1\;2\;3\;4)\,(5\;6\;7\;8\;9)\cdot(0\;1\;5\;t\;e)=(1\;2\;3\;4)\,(5\;6\;7\;8\;9\;t\;e\;0);$$

$$(0\;1\;2\;3\;4)\,(5\;6\;7\;8\;9)\cdot(0\;3\;5\;t\;e)=(3\;4)\,(5\;6\;7\;8\;9\;t\;e\;0\;1\;2).$$

To give the flavor of the arguments needed to establish the above, we sketch the proof of (ii) when $k=3$.

First, if n is odd $\geqq 9$, and if $m \geqq 3$, the product $\alpha\beta$ is an n-cycle. Here,

$$\alpha = (123)(456)\cdots(3m-2, 3m-1, 3m), \quad \text{and}$$

$$\beta = (34)(67)\cdots(3m-3, 3m-2) \equiv \beta_0, \quad \text{if} \quad n = 3m;$$

$$\beta = (n-1, n)\beta_0, \quad \text{if} \quad n = 3m+1;$$

$$\beta = (n-3, n-1)(n-2, n)\beta_0, \quad \text{if} \quad n = 3m+2.$$

The cases $5 \leqq n < 9$ are handled directly.

If the canonical factorization of P has only cycles with an odd number of symbols, there is nothing more to do. The other cases are handled by appropriate combinatorial arguments (and formulas).

The inductive argument used in proving (iii) is interesting, in that it proves more: the factors P_1, P_2 in the decomposition $P = P_1P_2$ can always be chosen so that the number of 3-cycles in P_1 differs by no more than 1 unit from the number of 3-cycles in P_2.

(v) If $n > 2$, every permutation in $\mathcal{A}_n$ is the product of two permutations of shape $2^1(n-2)^1$ [4]; the two factors lie in $\mathcal{A}_n$ if n is even.

Many similar results have since been found (e.g., [1]); we view these as having only tangential relevance; the PROBLEM does not require that the two factors have the same shape.

Burnside has a formula [3] for multiplication of conjugacy classes C_i:

$$C_iC_j = \Sigma_k c_{ijk} C_k.$$

For small values of n, the known result

$$c_{klm} = 0 \Leftrightarrow \Sigma_\chi \chi_k \chi_l \bar{\chi}_m / \chi_1 = 0$$

can be used to make verifications. (The summation Σ_χ runs over all irreducible characters; χ_k is the value of χ on the k-th class.)

Acknowledgement: NRC (Canada) Grant A5208.

References

1. E. A. Bertram, Even permutations as a product of two conjugate cycles, J. Combinatorial Theory, Ser. A, 12 (1972) 368–380.

2. J. L. Brenner, Problem E 1118, this MONTHLY, 62 (1955) 43.

3. W. Burnside, Theory of Groups of Finite Order, Cambridge University Press, New York, 1898.

4. Cheng-Hao Xu, The commutators of the alternating group, Sci. Sinica, 14 (1965) 339–342.

10 PHILLIPS ROAD, PALO ALTO, CA 94303, AND DEPARTMENT OF MATHEMATICS, UNIVERSITY OF VICTORIA, VICTORIA, B. C. V8W 2Y2, CANADA.

CLASSROOM NOTES

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Laboratoire Calcul des Probabilités, Université de Paris, T.56, 4 Place Jussieu, 75–230 Paris, France.

THEORY VS. MECHANICS IN AN APPLICATION OF CALCULUS TO BIOLOGY

ROCHELLE W. MEYER

Often students in calculus classes seem to view the underlying theory as something done by the instructor to delay presentation of the mechanics. Here is an example of an application of calculus to

biology in which a straightforward mechanical approach leads to a difficult situation, beyond most first year courses, but the theory of beginning calculus leads quickly to a solution.

Suppose a simple organism is being grown for food or medicinal purposes, e.g. yeast, and we desire to determine the amount and timing of periodic harvesting in order to maximize the yield per unit time.

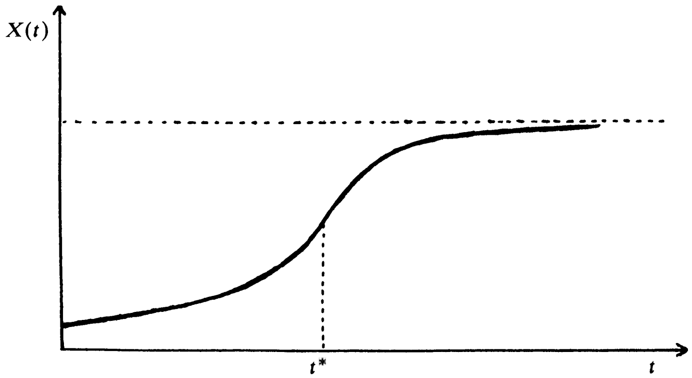


FIG. 1

Let $X(t)$ be a function giving the amount of organism present in a given culture at time t, assuming that no harvesting takes place. $X(t)$ describes the growth pattern of the organism; the specific parameters of such a function vary with such factors as nutrient medium, temperature, and the number of organisms at time $t = 0$. Functions like $X(t)$, whether determined empirically from the biological data or given in the explicit form of an equation modeling that data, often have a characteristic S-shaped graph, as shown in Figure 1. These S-shaped curves share the following properties:

1. *The curve is monotone strictly increasing, i.e., the first derivative is positive.*
2. *As $t \to +\infty$, the curve approaches a horizontal asymptote, the maximum carrying capacity of the culture environment.*
3. *The curve has an inflection point a t^*; the second derivative is positive for $t \in [0, t^*)$ and negative for $t \in (t^*, +\infty)$.*

One function which has these characteristics is the logistic, which is usually defined via a differential equation:

$$X'(t) = (r - \lambda X(t))X(t), \tag{1}$$

where r is the growth rate of the organism and λ reflects the carrying capacity of the environment.

For practical reasons, we plan to harvest periodically with periods of equal length and harvests of equal size. Some reflection shows that it is important to relate these fixed quantities so that the time between harvests be just enough for the culture to replace the amount harvested. Suppose we harvest an amount A at time t_h. Immediately thereafter, i.e., at the start of the next period, we have $X(t_h) - A$ as the amount of culture remaining. Since X is increasing there is exactly one number t_s such that $X(t_s) = X(t_h) - A$. We note that $t_s < t_h$ because the amount of culture after harvest is less than before harvest. Thus if we harvest the amount $A = X(t_h) - X(t_s)$ and wait $t_h - t_s$ between harvests, then harvest size and period length are properly related as required above. We define a yield-per-unit-time function $Y(t_s, t_h)$ as

$$Y(t_s, t_h) = \frac{X(t_h) - X(t_s)}{t_h - t_s}, \qquad t_s < t_h. \tag{2}$$

We wish to maximize $Y(t_s, t_h)$. A mechanical two-variable maximization is beyond the reach of a first year calculus course. If we assume t_s to be fixed we can try to maximize Y regarded as function of the single variable t_h. This approach leads easily to

$$(t_h - t_s)X'(t_h) = X(t_h) - X(t_s). \tag{3}$$

In order to finish solving for t_h we need explicit forms for X and X' to substitute into equation (3). In the case of the logistic we can get an explicit form for X' from equation (1) but an explicit form for X requires solving the differential equation, beyond the scope of most first year calculus courses. And if $X(t)$ is given as a set of data, not in equation form, we are even further from a solution. There must be a better way!

If we look again at the yield-per-unit-time function $Y(t_s, t_h)$ we note that the right hand side of that equation is a difference quotient for $X(t)$ involving two points $(t_s, X(t_s))$ and $(t_h, X(t_h))$, neither of which is fixed. $Y(t_s, t_h)$ gives the slope of the secant line intersecting the graph in those two points. The problem of maximizing $Y(t_s, t_h)$ can be translated to one of finding the secant with the greatest slope.

The Mean Value Theorem guarantees that there is a $t_m \in (t_s, t_h)$ such that the slope of the secant line is equal to the slope of the tangent at t_m. Since the derivative attains its maximum at t^*, if both t_s and t_h are on the same side of t^*, $X'(t_m) < X'(t^*)$. Therefore we want $t_s < t^* < t_h$.

Let L be the line tangent to the curve $X(t)$ at t^*. Because $X'(t)$ decreases for $t > t^*$, we have L above the curve for $t > t^*$. Similarly, L is below the curve for $t < t^*$. Therefore, for $t_s < t^* < t_h$, the secant line containing $(t_s, X(t_s))$ and $(t_h, X(t_h))$ must intersect L and have smaller slope than L. However, as $t_h - t_s$ becomes small, with $t_s < t^* < t_h$, the slope of the secant line approaches the slope of L as a limit. There is no pair (t_s, t_h) which maximizes the yield function $Y(t_s, t_h)$, thereby achieving maximum sustained harvest. But we can get arbitrarily close to the least upper bound of Y by harvesting small amounts near the inflection point of the curve X. Moreover, the least upper bound of Y is just $X'(t^*)$.

By appealing to calculus theory, we have solved the problem for the general S-shaped growth curve, without making use of the logistic. This means that it is not necessary to determine whether a specific set of data is best modeled by a logistic or to find the parameters of such a model. Simply plotting the data will suffice to indicate the neighborhood of the inflection point. If a particular logistic is known to model the growth, the inflection point can be easily located without solving the differential equation. The right hand side of equation (1), regarded as a function of X, has a graph which is an inverted parabola with roots at $X = 0$ and $X = r/\lambda$ and maximum point at $X = r/2\lambda$, midway between. Thus $X'(t)$ has its maximum for that value of t which gives $X(t) = r/2\lambda$.

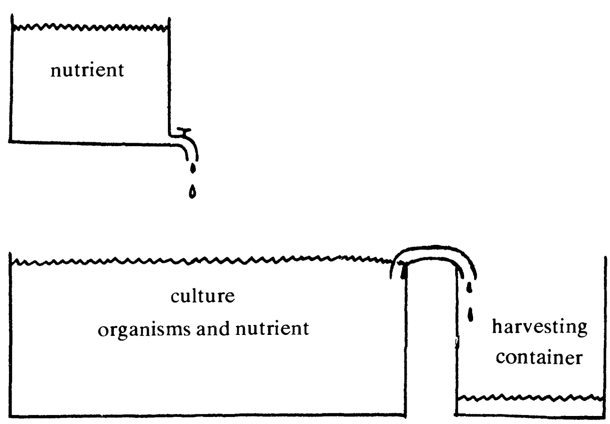


FIG. 2

This harvest plan is one of several used in laboratories. Other plans recommend themselves if we are concerned about condition of the organism (determined by position on the growth curve) or if we wish instead to maximize yield per unit of nutrient used. One apparatus used to grow and harvest simple organisms in laboratories is pictured in Figure 2. Each drop of nutrient added to the culture causes one drop of the culture, a mixture of nutrient and organisms, to overflow and be harvested. The size of each harvest is determined by the volume of the drop, the timing by the time between addition of drops. With this apparatus our theoretical solution of harvesting small amounts near the inflection point can be realized.

The author wishes to thank the referees for their helpful suggestions.

DEPARTMENT OF MATHEMATICS, SUFFOLK COUNTY COMMUNITY COLLEGE, SELDEN, NY 11784.

ON DECIDING WHETHER A SURFACE IS PARABOLIC OR HYPERBOLIC

JOHN MILNOR

Let M be a simply connected open surface, differentiable of class say C^∞, with a C^∞ Riemannian metric. It follows from complex function theory that M is conformally diffeomorphic either to the complex plane, or to the open unit disk in the complex plane. (See for example [Springer, p. 225] together with the local argument as given in [Chern] or [Chern, Hartman, Wintner] or [Courant–Hilbert, p. 350] or [Milnor, pp. 1110, 1111].) In the first case one says that M is **parabolic**, in the second case **hyperbolic**.

This note will attack the problem of deciding whether M is parabolic or hyperbolic in the very special circumstance that M is complete and rotationally symmetric about a point p, so that the Gaussian curvature K can be expressed as a function of the distance r from p. Let $2\pi g(r)$ denote the circumference of the circle of radius r centered at p. Introducing geodesic polar coordinates r, θ the Riemannian metric takes the form $dr^2 + g(r)^2 d\theta^2$. Computation shows that Gauss curvature is given by

$$K = -(d^2g/dr^2)/g$$

or briefly, $K = -\ddot{g}/g$. (See for example [**18**, pp. 136–138] or [**16**, pp. 3B, 32–37].) Thus, the function g satisfies the linear differential equation $\ddot{g} + Kg = 0$ with initial conditions $g(0) = 0$, $\dot{g}(0) = 1$.

THEOREM. *If $K \geqq -1/(r^2 \log r)$ for large r, then the surface is conformally parabolic. But if $K \leqq -(1+\varepsilon)/(r^2 \log r)$ for large r, and if the function $g(r)$ is unbounded, then the surface is hyperbolic.*

Here $\varepsilon > 0$ can be any positive constant.

Note that the hypothesis that $g(r)$ is unbounded is essential. In fact, if $g(r)$ is bounded, then the surface is necessarily parabolic. Compare the lemma below. An example is provided by the function $g(r) = re^{-r^2}$, with $K(r) = 6 - 4r^2$. This yields a parabolic surface since g is bounded, even though K tends to $-\infty$ as $r \to \infty$. (Such an example, with g bounded, cannot have $K \leqq 0$ for all values of r. For if $K \leqq 0$ or equivalently $\ddot{g} \geqq 0$ for all r, then it follows by integration that $\dot{g} \geqq 1$ and $g(r) \geqq r$ so that g must be unbounded.)

This theorem surprised the author: It seemed strange that a parabolic surface could be converted into a hyperbolic surface simply by replacing its curvature function $K(r) = -1/(r^2 \log r)$ for $r \geqq c$ by $(1+\varepsilon)K(r)$ for $r \geqq c$.

The proof will be based on the following. (Compare [**1**].) Choose $a > 0$.

LEMMA. *The surface M is hyperbolic if and only if the integral $\int_a^\infty dr/g(r)$ is finite.*

Proof. Introduce a new coordinate ρ in place of r by setting $\rho = \int_a^r ds/g(s)$. Then the metric takes

the form

$$dr^2 + g^2 d\theta^2 = g^2(d\rho^2 + d\theta^2).$$

(Coordinates ρ, θ for which the metric takes the latter form, conformally equivalent to the flat metric $d\rho^2 + d\theta^2$, are known classically as *isothermal coordinates.*) Now map M conformally into the complex numbers by sending the point with coordinates ρ, θ to the complex number $\exp(\rho + i\theta) = e^\rho \cos\theta + ie^\rho \sin\theta$. This conformal mapping is at first not defined at the base point p. However, it is bounded, smooth, and conformal throughout a neighborhood of p, except at the point p itself. Hence the apparent singularity at p is *removable*, and we obtain a smooth conformal mapping which is defined and one-to-one throughout M.

[Expressing this construction in a different way, we map the point of M with geodesic polar coordinates (r, θ) to the complex number $Re^{i\theta}$ with polar coordinates (R, θ), where $R = \exp(\int_a^r ds/g(s)) = \exp(\int_1^r ds/g(s) + \text{constant})$. This last formula for R is the most general one which yields a radial stretching factor dR/dr precisely equal to the circumferential stretching factor $2\pi R/2\pi g(r)$, so that the mapping is conformal.]

Evidently the image of this mapping is either the disk of radius $\exp \int_a^\infty dr/g(r)$ or the entire complex plane according as this integral is finite or infinite. ■

Remark. If the total absolute curvature $\int\int |K| dA$ of a complete simply connected surface is finite, then [Blanc and Fiala] have shown that the surface is parabolic. Compare **[8]**, **[6]**. (In the rotationally symmetric case we can recover this result from the Lemma as follows. Since the Riemannian area element dA is equal to $gdrd\theta$, it follows that

$$(2\pi)^{-1} \int\int |K| dA = \int_0^\infty |K| g dr = \int_0^\infty |\ddot{g}| dr$$

is equal to the total variation of $\dot{g}$. If this total variation is finite, then $\dot{g}$ is bounded, hence $g(r) \leqq Cr$, and $\int_1^\infty dr/g = \infty$, so that the surface must be parabolic.)

Here is a basic example which illustrates the Lemma. Consider the function $g_0(r) = r \log r$ for $r \geqq 2$ (choosing say $g_0(r) = r$ for $r \leqq 1$ and interpolating smoothly in the interval $1 \leqq r \leqq 2$). The corresponding surface is parabolic since $\int_2^\infty dr/(r \log r) = [\log\log r]_2^\infty = \infty$. The associated Gauss curvature K_0 is given by

$$K_0 = -\ddot{g}_0/g_0 = -1/(r^2 \log r)$$

for $r \geqq 2$. (In this example, the total absolute curvature $\int\int |K_0| dA_0$ is infinite, so that the Blanc–Fiala criterion does not apply.)

Proof of Theorem. Let $g(r) > 0$ be any smooth function whose associated curvature $K = -\ddot{g}/g$ satisfies the inequality $K(r) \geqq K_0(r)$ for $r \geqq a$; where $K_0 = -1/(r^2 \log r)$ as above. (Note that K may take both positive and negative values.) Multiplying g by a small positive constant if necessary, we may assume that the initial inequalities

$$g(a) < g_0(a), \tag{1}$$

$$\dot{g}(a) < \dot{g}_0(a) \tag{2}$$

are satisfied. *It follows that* $g(r) < g_0(r)$ *for all* $r \geqq a$. For otherwise, if b were the smallest number greater than a with $g(b) = g_0(b)$, then integrating the inequality

$$\ddot{g} = -Kg \leqq -K_0 g_0 = \ddot{g}_0$$

from a to r and adding (2) we would obtain $\dot{g}(r) < \dot{g}_0(r)$ for $a \leqq r \leqq b$. Integrating this new inequality from a to b and adding (1) we would obtain $g(b) < g_0(b)$, which contradicts the choice of b. Thus $g(r) < g_0(r)$ for all $r \geqq a$. Hence the integral $\int_a^\infty dr/g(r)$ is also divergent, and the surface is parabolic.

For the proof in the other direction, we use a comparison function of the form

$$g_\varepsilon(r) = r(\log r)^{1+\varepsilon} \text{ for } \quad r \geqq 2.$$

This yields a hyperbolic surface since the integral

$$\int_2^\infty dr/r(\log r)^{1+\varepsilon} = [-1/\varepsilon(\log r)^\varepsilon]_2^\infty$$

is finite. Computation shows that the corresponding Gauss curvature is given by

$$K_\varepsilon = -(1+\varepsilon)(1+\varepsilon/\log r)/(r^2 \log r) \geqq -(1+2\varepsilon)/(r^2 \log r)$$

for large r.

Now consider a smooth function $g(r) > 0$ whose associated Gauss curvature $K = -\ddot{g}/g$ satisfies

$$K(r) \leqq -(1+2\varepsilon)/(r^2 \log r) \leqq K_\varepsilon(r)$$

for $r \geqq a$. If $\dot{g}(a) > 0$ then, after multiplying the function $g(r)$ by a large constant if necessary, a completely analogous argument shows that $g(r) > g_\varepsilon(r)$ for all $r \geqq a$. Hence the integral $\int_a^\infty dr/g(r)$ is also convergent, and the surface is hyperbolic.

This argument will work whenever $\dot{g}(r) > 0$ for *some* sufficiently large r. In particular, it will certainly work whenever the function $g(r)$ is unbounded. On the other hand, if $\dot{g}(r) \leqq 0$ for all large r, then the function $g(r)$ is bounded, and hence by the Lemma the surface is parabolic. ■

In conclusion, here are two further problems:

(1) *Does the Theorem proved above remain valid in the more general case of a complete Riemannian surface which is not rotationally symmetric?*

The statement would still make sense as long as the base point p has no conjugate points, so that there exists a global geodesic polar coordinate system with metric of the form $dr^2 + g(r,\theta)^2 d\theta^2$. Robert Osserman points out to the author that if $K(r,\theta) \geqq -1/(r^2 \log r)$ for large r then, using a theorem of Ahlfors [1] in place of the Lemma, one can indeed prove that M is parabolic. Greene and Wu [7] prove an analogous result in higher dimensions, assuming the sharper hypothesis that all sectional curvatures satisfy $0 \geqq K \geqq -C/r^{2+\varepsilon}$. For discussion of a related problem in the hyperbolic case, (see [12] §6.4).

(2) Consider an embedded surface in Euclidean 3-space of the form $z = f(x, y)$. *How can one decide effectively whether this surface is parabolic or hyperbolic?* This problem has been studied by Osserman [**12, 13**], H. Huber [**9**], Jenkins [**10**], and by Rüedy [**14**].

References

1. L. Ahlfors, Sur le type d'une surface de Riemann, C. R. Acad. Sci., Paris, 201 (1935) 30–32. (Compare [Sario and Nakai, p. 329].)

2. C. Blanc and F. Fiala, Le type d'une surface et sa courbure totale, Comment. Math. Helv., 14 (1941–42) 230–233.

3. S. S. Chern, An elementary proof of the existence of isothermal parameters on a surface, Proc. Amer. Math. Soc., 6 (1955) 771–782.

4. S. S. Chern, P. Hartman, and A. Wintner, On isothermic coordinates, Comment. Math. Helv., 28 (1954) 301–309.

5. R. Courant and D. Hilbert, Methods of Mathematical Physics, vol. 2, Interscience, Wiley, New York, 1962.

6. R. Finn, On a class of conformal metrics, with application to differential geometry in the large, Comment. Math. Helv., 40 (1965–66) 1–30.

7. R. E. Greene and H. Wu, Curvature and complex analysis, II, Bull. Amer. Math. Soc., 78 (1972) 866–870.

8. A. Huber, On subharmonic functions and differential geometry in the large, Comment. Math. Helv., 32 (1957) 13–72.

9. H. Huber, Riemannsche Flächen von hyperbolischem Typus im euklidischen Raum, Math. Ann., 139 (1959) 140–146.

10. J. A. Jenkins, On hyperbolic surfaces in three-dimensional euclidean space, Michigan Math. J., 8 (1961) 1–5.
11. J. Milnor, A problem in cartography, this MONTHLY, 76 (1969) 1101–1112.
12. R. Osserman, Riemann surfaces of class A, Trans. Amer. Math. Soc., 82 (1956) 217–245.
13. ———, Hyperbolic surfaces of the form $z = f(x, y)$, Math. Ann., 144 (1961) 77–79.
14. R. Rüedy, Embeddings of open Riemann surfaces, Comment. Math. Helv., 46 (1971) 214–225.
15. L. Sario and M. Nakai, Classification Theory of Riemann Surfaces, Grundlehren 164, Springer-Verlag, New York, 1970.
16. M. Spivak, A Comprehensive Introduction to Differential Geometry II, Publish or Perish, Inc., Boston, 1970.
17. G. Springer, Introduction to Riemann Surfaces, Addison–Wesley, Reading, Mass., 1957.
18. D. Struik, Lectures on Classical Differential Geometry, Addison–Wesley, Reading, Mass., 1950.

INSTITUTE FOR ADVANCED STUDY, PRINCETON, N.J. NJ 08540.

CYCLIC SUBGROUPS OF THE PRIME RESIDUE GROUP

F. RUDOLF BEYL

The study of metacyclic groups, i.e., groups which are extensions of a cyclic group by a cyclic group, leads immediately to the following problem of elementary number theory: *Which prime residues* (mod M, *say*) *are mutual powers of each other*? Here the term prime residue mod M means a residue class modulo M which can be represented by an integer prime to M.

If we shift attention from the prime residue to the cyclic subgroup it generates in the group $\mathbb{Z}_M^*$ of all prime residues mod M, the problem reads: Describe the cyclic subgroups of $\mathbb{Z}_M^*$. The following can be treated in a course on elementary number theory just before primitive elements. Most of the statements can be found in Basmaji [**1**, § 3 p. 177] as well as in other papers on metacyclic groups, but the presentations there are not fit for classroom use.

If n is a natural number and p a prime, let $\#_p(n)$ denote the exact power of p dividing n.

PROPOSITION 1: *Let p be a prime, r an integer* $\equiv 1$ mod p, $r \neq -1$, *let n be a natural number and* $S = 1 + r + r^2 + \cdots + r^{n-1}$. *If p is odd, then* $\#_p(S) = \#_p(n)$. *If $p = 2$ and n is odd, then* $\#_2(S) = \#_2(n) = 0$. *If $p = 2$ and n is even, then* $\#_2(S) = \#_2(n) + \#_2(r+1) - 1$, *so that* $\#_2(S) = \#_2(n)$ *if* $r \equiv 1$ mod 4.

PROPOSITION 2: *Let p be a prime,* $m \geqq 1$, *and* $M = p^m$. *The prime residue r lies in the p-Sylow subgroup of* $\mathbb{Z}_M^*$ *if and only if* $p \mid r - 1$. *Let p^q be the order of such an r.*

CASE (i): *p odd. Then* $q = \max(m - \#_p(r-1), 0)$.

CASE (ii): $p = 2$. *If* $r \equiv 1$ *then* $q = 0$, *and if* $r \equiv -1$ *then* $q = 1$. *For* $m \geqq 3$ *any other prime residue has the form* $r = 4w + 1$ *or* $r = 4w - 1$ *with* $2^{m-2} \nmid w$; *in either case* $q = m - 2 - \#_2(w)$.

Proof of Proposition 1, by induction on the number of prime factors in n. The case $n = 1$ is trivial. Next let n be a prime. If $n \neq p$ then $S \equiv n$ mod p, and we are done. If $p = n = 2$, then $S = r + 1$, and we are also done. If $p = n$ is odd, then

$$S = \sum (1 + \lambda p)^i \equiv p + \frac{1}{2} \cdot p \cdot (p-1) \cdot \lambda p \equiv p \mod p^2,$$

as desired. Now consider composite n. Let $n = k \cdot l$ with $1 < k < n$, and assume $2 | k$ if $2 | n$. Then

$$S = \left(\sum_{i=0}^{k-1} r^i\right) \cdot \left(\sum_{i=0}^{l-1} s^i\right), \qquad s = r^k. \tag{1}$$

$\mathbb{Z}_m^* \times \mathbb{Z}_n^*$ when $M = m \cdot n$ with $(m, n) = 1$. If m is an odd prime power, then $\mathbb{Z}_m^*$ is cyclic (existence of a primitive element) and the subgroups of $\mathbb{Z}_m^*$ are completely characterized by their order. If m is a power of 2, then the cyclic subgroups can be determined by Prop. 4. Thus the following proposition gives, at least in principle, the answer to our problem by iteration.

PROPOSITION 5: *Let $G = A \times B$ be a direct product of groups. If the elements $r = (a_r, b_r)$ and $s = (a_s, b_s)$ generate the same cyclic subgroup in G, then so do a_r and a_s in A and b_r and b_s in B. Conversely, given $r = (a_r, b_r)$ and $s = (a_s, b_s)$ with $a_s = a_r^\alpha$ and $b_s = b_r^\beta$ for some α prime to $t =$* order (a_r) *and some β prime to $\tau =$* order (b_r); *then r and s generate the same cyclic subgroup in G if, and only if, $(t, \tau) \mid \beta - \alpha$.*

Proof. Assume the hypothesis of the last assertion. We claim that the properties (i) to (iv) in the following chain are mutually equivalent: (i) there is an integer γ with $s = r^\gamma$; (ii) there is γ with $\alpha \equiv \gamma \bmod t$ and $\beta \equiv \gamma \bmod \tau$; (iii) the diophantine equation $\lambda t - \mu\tau = \beta - \alpha$ has solutions λ and μ; (iv) $(t, \tau) \mid \beta - \alpha$. Indeed, when (ii) is given, let $\lambda = (\gamma - \alpha)/t$ and $\mu = (\gamma - \beta)/\tau$ with $\lambda t - \mu\tau = \beta - \alpha$. Given (iii), let $\gamma = \alpha + \lambda t = \beta + \mu\tau$. Thus (ii) and (iii) are equivalent. The equivalences (i)$\leftrightarrow$(ii) and (iii)$\leftrightarrow$(iv) are clear. □

NOTE. The following variant of Proposition 1 is also useful and was pointed out to us by A. Brandis.

PROPOSITION 1′: *In the hypotheses of Proposition* 1, *let p be an odd prime and $n = p^m$ for $m \geqq 0$. Then $S \equiv p^m \bmod p^{m+1}$.*

Proof of Proposition 1′, by induction on m. Use $S \equiv p \bmod p^2$ for $m = 1$ from our proof of Prop. 1. Do the induction step ($m \geqq 2$) with Formula (1), where $n = p^m = k \cdot l$ and $k \neq 1 \neq l$. □

References

1. B. G. Basmaji, On the isomorphisms of two metacyclic groups, Proc. Amer. Math. Soc., 22 (1969) 175–182.
2. C. Chevalley, Sur certains groupes simple, Tôhoku Math. J. (2), 7 (1955) 14–66.

MATHEMATISCHES INSTITUT DER UNIVERSITÄT, 69 HEIDELBERG 1, IM NEUENHEIMER FELD 288, W. GERMANY.

PROOF BY MECHANICS THAT THE ALTITUDES OF A TETRAHEDRON ARE ASSOCIATED

LEON GERBER

It is well known that while the altitudes of a triangle concur, the altitudes of a tetrahedron do not concur in general. Using only simple results from elementary mechanics we shall fulfill the promise of the title by showing directly that every line which meets three of the altitudes of a tetrahedron also meets the fourth; and further, that if any two altitudes meet, so do the other two. (See **[1–4]** for extensions of these results to n dimensions.)

Although the possibility has long been denied by teachers of elementary geometry, our proof is the extension to three dimensions of the following familiar proof that the altitudes of a triangle $\mathcal{A}$ are concurrent.

Let $\mathcal{B}$ be the anticomplementary triangle of $\mathcal{A}$, that is, its image in the dilation $D(G, -2)$ centered at the centroid G of $\mathcal{A}$ and having constant -2. The vertices of $\mathcal{A}$ are the midpoints of the sides of $\mathcal{B}$, so the altitudes of $\mathcal{A}$ are the perpendicular bisectors of the sides of $\mathcal{B}$. These perpendicular bisectors concur at the circumcenter of $\mathcal{B}$, which is then the orthocenter of $\mathcal{A}$. (Since the circumcenter of $\mathcal{B}$ is

the image of the circumcenter of $\mathscr{A}$ in the dilation, this proves the additional fact that the centroid divides the segment joining the orthocenter and circumcenter in the ratio 2:1.)

Passing to three dimensions, let $\mathscr{A}$ be a tetrahedron with centroid G and let $\mathscr{B}$ be its image in $D(G,-3)$, so the vertices of $\mathscr{A}$ are the centroids of the faces of $\mathscr{B}$ and the altitudes of $\mathscr{A}$ are the perpendiculars to the faces of $\mathscr{B}$ at their centroids. There is no way on earth that we can prove these perpendiculars are concurrent, so let us put $\mathscr{B}$ in a pressurized container and take it out of earth's gravitational field. Then $\mathscr{B}$ will be in equilibrium under the forces due to uniform pressure on the faces of $\mathscr{B}$. It is an elementary fact of mechanics that the uniform pressure on each face of the tetrahedron can be replaced by a single force acting at its centroid and normal to the face, that is, along an altitude of $\mathscr{A}$.

Recall that the moment of a nonzero force about a line is the product of the projection of the force on a plane perpendicular to the line, and the distance from the line of action of the force to the line, and vanishes if and only if the two lines are coplanar. Since $\mathscr{B}$ is in equilibrium, the moments of these four forces about any line sum to zero. If we consider a line which meets three of the altitudes, the three corresponding moments are zero and hence the fourth moment is also zero. Since the force is not zero, the line must also meet the fourth altitude. This concludes the proof.

Suppose two of the altitudes meet at the point P. Taking moments about P we see that the sum of the moments of the other two forces is zero. Thus they are coplanar, i.e., the other two altitudes also intersect and their plane contains P.

References

1. L. Gerber, The altitudes of a simplex are associated, Math. Mag., 46 (1973) 155–157. MR 47 # 7580.
2. ———, Associated and perspective simplexes, Trans. Amer. Math. Soc., 201 (1975) 43–55. MR 50 # 8262.
3. ———, Associated and skew-orthologic simplexes, ibid. to appear.
4. S. R. Mandan, Uni- and demi-orthocentric simplexes, J. Indian Math. Soc., (NS) 23 (1959) 169–184, ibid. 26 (1962) 5–11. MR 26 # 2922.

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, ST. JOHN'S UNIVERSITY, JAMAICA, NY 11439.

ANOTHER NOTE ON LIMITS OF COMPOSITE FUNCTIONS

JÜRG RÄTZ

1. For the evaluation of

$$\lim_{x\to 0}\left(1-\frac{\sin x}{x}\right)^{-1},$$

a very natural first approach using no derivatives runs as follows: As x tends to 0, $(\sin x/x)$ tends to 1 from below, thus the required limit exists and is $+\infty$. This example and, of course, many others, might illustrate that it is desirable to have at hand a sufficiently flexible theorem on limits of composite functions which turns the argument used above into a rigorous method. The purpose of this note is to present such a theorem; in a more special version, it has been a subject of the author's undergraduate analysis course for several years.

There are good reasons for discussing the question on the level of functions between topological spaces. The corollaries for the familiar spaces of classical analysis will be immediate.

2. By a *filter* on a nonempty set X we mean here a nonempty collection $\boldsymbol{F}$ of subsets of X with the properties $A, B \in \boldsymbol{F} \Rightarrow A \cap B \in \boldsymbol{F}$ and $A \in \boldsymbol{F}, A \subset B \subset X \Rightarrow B \in \boldsymbol{F}$ ([**4**], p. 29). Most texts require that the empty set not be a member of a filter. For our purposes, however, the present definition is more appropriate: As we do not insist on the uniqueness of limits, it is quite natural to consider also

limits at isolated points, and the main results remain valid under our more general definition. $\boldsymbol{F}$ is called *proper* if $\varnothing \notin \boldsymbol{F}$. If $\boldsymbol{F}$ is a filter on X and X_0 a nonempty subset of X, then $\boldsymbol{F}_0 = \{F \cap X_0 \mid F \in \boldsymbol{F}\}$ is a filter on X_0, the so-called *trace* of $\boldsymbol{F}$ in X_0. Let $\boldsymbol{B}$ be a nonempty collection of subsets of X such that for any $B_1, B_2 \in \boldsymbol{B}$ there exists $B_3 \in \boldsymbol{B}$ with the property $B_3 \subset B_1 \cap B_2$. Then $\boldsymbol{B}$ is called a *filter base* on X, and the collection $\{F \subset X \mid$ there exists $B \in \boldsymbol{B}$ such that $B \subset F\}$ then is a filter on X; it is said to be *generated* by $\boldsymbol{B}$.

If $\boldsymbol{F}$ is a filter on X and $f: X \to Y$, then the filter on Y generated by the filter base $\{f(A) \mid A \in \boldsymbol{F}\}$ is denoted by $f(\boldsymbol{F})$. For filters $\boldsymbol{F}$ on X, $\boldsymbol{G}$ on Y, the statement $\boldsymbol{G} \subset f(\boldsymbol{F})$ is equivalent to the fact that for any $G \in \boldsymbol{G}$ there exists $F \in \boldsymbol{F}$ such that $f(F) \subset G$. An essential part of what we are considering here is based upon the following set-theoretical lemma the proof of which is left to the reader.

LEMMA. *Let $\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{H}$ be filters on the nonempty sets X, Y, Z, respectively, and $f: X \to Y$, $g: Y \to Z$. If $\boldsymbol{G} \subset f(\boldsymbol{F})$ and $\boldsymbol{H} \subset g(\boldsymbol{G})$, then also $\boldsymbol{H} \subset (g \circ f)(\boldsymbol{F})$.*

COROLLARY. *Let X, Y, Z be topological spaces, $f: X \to Y$, $g: Y \to Z$, and $x \in X$. If f is continuous at x and g is continuous at $f(x)$, then $g \circ f$ is continuous at x.*

Proof: Let $\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{H}$ be the neighborhood filters $\boldsymbol{U}(x)$, $\boldsymbol{U}[f(x)]$, $\boldsymbol{U}[(g \circ f)(x)]$, respectively, and notice that f is continuous at x iff $\boldsymbol{U}[f(x)] \subset f[\boldsymbol{U}(x)]$, etc. Now apply the lemma.

REMARK 1. The lemma yields a full analogue to the corollary for uniformly continuous functions between uniform spaces (see, e.g., [**7**], p. 175, Definition 20.3).

3. For the attack of the main result, we need more prerequisites. The basic idea is to consider, besides the neighborhood filters, also other *open filters*, i.e., filters generated by open sets. This amounts to adjoining ideal points to the given topological space. For the extended real and complex number systems cf., e.g., [**1**], p. 56–57; for general topological spaces [**4**], p. 158–161, or [**7**], p. 131 ff. But we emphasize that our primary interest is in the neighborhoods of these ideal points rather than in the points themselves. Thus we only borrow the language and formalism from a standard extension procedure in general topology without really performing the procedure. In this way, the character of the discussion in kept elementary. A more serious reason, however, is mentioned in Remark 7 below.

Let X be a topological space and $\Omega(X)$ the collection of all open filters on X. Then, for any $x \in X$, the neighborhood filter $\boldsymbol{U}(x)$ belongs to $\Omega(X)$. If X is the real line $\mathbb{R}$, also the filters $\boldsymbol{U}_+(x)$, $\boldsymbol{U}(x)$, $\boldsymbol{U}(+\infty)$, and $\boldsymbol{U}(-\infty)$ generated by the interval systems $\{(x, x+\varepsilon) \mid \varepsilon > 0\}$, $\{(x-\varepsilon, x) \mid \varepsilon > 0\}$, $\{(1/\varepsilon, \rightarrow) \mid \varepsilon > 0\}$, and $\{(\leftarrow, -1/\varepsilon) \mid \varepsilon > 0\}$, respectively, are members of $\Omega(X)$. The role they play here is obvious.

Let be $\hat{X} = X \cup \{\boldsymbol{F} \in \Omega(X) \mid \boldsymbol{F} \neq \boldsymbol{U}(x)$ for every $x \in X\}$, and let $\sigma: \hat{X} \to \Omega(X)$ be defined by $\sigma(\hat{x}) = \boldsymbol{U}(\hat{x})$ for $\hat{x} \in X$; $\sigma(\hat{x}) = \hat{x}$ for $\hat{x} \in \hat{X} \backslash X$, i.e., for the ideal points of X. (By the way, σ is bijective iff X is a T_0-space.)

Following the pattern of definition of limits of functions in analysis, we need the auxiliary filters $\dot{\boldsymbol{U}}(x)$ generated by $\{U \backslash \{x\} \mid U \in \boldsymbol{U}(x)\}$, i.e., by the deleted neighborhoods of $x \in X$. (Note that $\dot{\boldsymbol{U}}(x) \in \Omega(X)$ if X is a T_1-space, but not in general.) For isolated points x of X, $\dot{\boldsymbol{U}}(x)$ is the improper filter $\boldsymbol{P}(X)$ consisting of all subsets of X. Finally, we define the operator ρ by $\rho(\hat{x}) = \dot{\boldsymbol{U}}(\hat{x})$ for $\hat{x} \in X$; $\rho(\hat{x}) = \hat{x}$ for $\hat{x} \in \hat{X} \backslash X$. Now we are prepared for a unified treatment of ordinary, improper, and — in the case of linearly ordered spaces — one-sided limits. For a complete settlement of the case of ordinary limits see [**6**]. A unified theory of limits using nets can be found in [**5**].

DEFINITION. *If X and Y are topological spaces and $f: X \to Y$, $\hat{x} \in \hat{X}$, $\hat{y} \in \hat{Y}$, then $\hat{y}$ is called a limit of f at $\hat{x}$, in symbols $\hat{y} \in \operatorname{Lim}(f; \hat{x})$, iff $\sigma(\hat{y}) \subset f[\rho(\hat{x})]$.*

THEOREM. *Let X, Y, Z be topological spaces; $f: X \to Y$, $g: Y \to Z$; $\hat{x} \in \hat{X}$, $\hat{y} \in \hat{Y}$, $\hat{z} \in \hat{Z}$, and $\hat{y} \in \operatorname{Lim}(f; \hat{x})$. Then we have:*

(a) *If $\hat{x}$ is an isolated point of X, then* $\operatorname{Lim}(g \circ f; \hat{x}) = \hat{Z}$.

(b) *If* $\hat{y} \in \hat{Y} \setminus Y$ *and* $\hat{z} \in \operatorname{Lim}(g; \hat{y})$, *then* $\hat{z} \in \operatorname{Lim}(g \circ f; \hat{x})$.

(c) *If* $\hat{y} \in Y$, $\hat{z} \in \operatorname{Lim}(g; \hat{y})$ *and if there exists* $U_1 \in \rho(\hat{x})$ *with the property* $\hat{y} \notin f(U_1)$, *then* $\hat{z} \in \operatorname{Lim}(g \circ f; \hat{x})$.

(d) *If* $\hat{y} \in Y$ *and if there exists* $U_0 \in \rho(\hat{x})$ *such that* $\hat{y} \in f(U_0)$ *and that the restricted function* $g \mid f(U_0)$ *is continuous at* $\hat{y}$, *then* $g(\hat{y}) \in \operatorname{Lim}(g \circ f; \hat{x})$.

Proof: (a) Since $\rho(\hat{x}) = \boldsymbol{P}(X)$, it follows that $(g \circ f)[\rho(\hat{x})] = \boldsymbol{P}(Z)$, i.e., that $\sigma(\hat{z}) \subset (g \circ f)[\rho(\hat{x})]$ for every $\hat{z} \in \hat{Z}$.

(b) We have $\sigma(\hat{y}) \subset f[\rho(\hat{x})]$, and from $\hat{y} \in \hat{Y} \setminus Y$ we obtain $\rho(\hat{y}) = \hat{y} = \sigma(\hat{y})$, hence $\rho(\hat{y}) \subset f[\rho(\hat{x})]$. This and $\sigma(\hat{z}) \subset g[\rho(\hat{y})]$ imply $\sigma(\hat{z}) \subset (g \circ f)[\rho(\hat{x})]$ by virtue of the lemma.

(c) Let $V \in \rho(\hat{y})$ be arbitrary. As $\hat{y} \in Y$, we get $V \cup \{\hat{y}\} \in \sigma(\hat{y})$. So there exists $U \in \rho(\hat{x})$ such that $f(U) \subset V \cup \{\hat{y}\}$. Now $U \cap U_1 \in \rho(\hat{x})$ and $f(U \cap U_1) \subset f(U) \cap f(U_1) \subset (V \cup \{\hat{y}\}) \cap (Y \setminus \{\hat{y}\}) \subset V$, therefore $V \in f[\rho(\hat{x})]$, i.e., $\rho(\hat{y}) \subset f[\rho(\hat{x})]$, and the proof is completed as in part (b).

(d) Let $W \in \sigma[g(\hat{y})]$ be arbitrary. Then there exists $V \in \sigma(\hat{y})$ such that $g(V \cap f(U_0)) \subset W$. There exists $U \in \rho(\hat{x})$ satisfying $f(U) \subset V$. Since $U \cap U_0 \in \rho(\hat{x})$ and $f(U \cap U_0) \subset f(U) \cap f(U_0) \subset V \cap f(U_0)$, we have $g[f(U \cap U_0)] \subset g(V \cap f(U_0)) \subset W$, i.e., $g(\hat{y}) \in \operatorname{Lim}(g \circ f; \hat{x})$. (As a matter of fact, this is again an application of the lemma: $\boldsymbol{G}$ is the trace of $\sigma(\hat{y})$ in $f(U_0)$ and $\boldsymbol{F}$ the trace of $\rho(\hat{x})$ in U_0.)

4. REMARK 2. For our initial example we put $X = \mathbb{R} \setminus \{0\}$, $Y = \mathbb{R} \setminus \{1\}$, $Z = \mathbb{R}$, each with the usual topology; $f(x) = \sin x / x$, $g(y) = (1-y)^{-1}$. Here $\hat{x}$, $\hat{y}$, $\hat{z}$ are ideal points of X, Y, Z, respectively: They are the traces in X, Y, Z of the filters $\dot{\boldsymbol{U}}(0)$, $\boldsymbol{U}_{-}(1)$, $\boldsymbol{U}(+\infty)$ defined earlier. Now part (b) of the theorem tells us that the result obtained by the natural approach is correct.

REMARK 3. The question of uniqueness of limits should not be neglected. First, in part (a) of the theorem, no separation axiom for Z can prevent $\operatorname{Lim}(g \circ f; \hat{x})$ from being so large. But also for non-isolated points $\hat{x}$ of X, $\operatorname{Lim}(f; \hat{x})$ need not be a singleton: Y can fail to be a T_2-space, and, in general, $\hat{Y}$ contains too many ideal points. It is easily seen from the definition that the following statement holds:

If $\hat{Y}_0$ *is a subset of* $\hat{Y}$ *such that for any distinct elements* $\hat{y}_1$, $\hat{y}_2$ *of* $\hat{Y}_0$, $\sigma(\hat{y}_1) \cup \sigma(\hat{y}_2)$ *generates the improper filter* $\boldsymbol{P}(Y)$ (*i.e., there exist* $U_1 \in \sigma(\hat{y}_1)$, $U_2 \in \sigma(\hat{y}_2)$ *with the property* $U_1 \cap U_2 = \varnothing$), *then* $\hat{Y}_0 \cap \operatorname{Lim}(f; \hat{x})$ *contains at most one element provided* $\rho(\hat{x}) \neq \boldsymbol{P}(X)$.

Notice that this applies to the extended real and complex number systems $\hat{\mathbb{R}}_0 = \mathbb{R} \cup \{+\infty, -\infty\}$ and $\hat{\boldsymbol{C}}_0 = \mathbb{C} \cup \{\infty\}$.

REMARK 4. In connection with part (d) of the theorem, the example $X = Y = Z = \mathbb{R}$, $f \equiv 0$, $g(y) = 0$ for $y \neq 0$, $g(0) = 1$ is quite famous ([**3**], p. 26). Part (d) is appropriate for unmasking and explaining this example and similar ones.

REMARK 5. The hypothesis in (d) does not imply the continuity of $g \mid f(X)$ at $\hat{y}$: Let be $X = Y = \mathbb{R}^2$, $\|\cdot\|$ the euclidean norm on $\mathbb{R}^2$, and Z the real line. $f\colon X \to Y$ and $g\colon Y \to Z$ are given by

$$f(x) = f(x_1, x_2) = \begin{cases} (x_1, 0) & \text{if } \|x\| < 1, \\ (0, x_2) & \text{if } \|x\| \geqq 1, \end{cases}$$

$$g(y) = g(y_1, y_2) = \begin{cases} 0 & \text{if } y_2 = 0, \\ 1 & \text{if } y_2 \neq 0. \end{cases}$$

Then $\operatorname{Lim}(f; (0,0)) = \{(0,0)\}$, and for $U_0 = \{x \in X \mid 0 < \|x\| < 1\}$ we get $f(U_0) = \{(x_1, 0) \mid -1 < x_1 < 1\}$. So $g \mid f(U_0)$ is continuous at $(0,0)$, but $g \mid f(X)$ is not. Notice, however, that if $f(U_0)$ is a neighborhood of $\hat{y}$ in $f(X)$, then the continuity of $g \mid f(X)$ at $\hat{y}$ follows.

REMARK 6. Part (d) of the theorem has no converse in the following sense: $\hat{y} \in f(U)$ for every $U \in \rho(\hat{x})$ and $g(\hat{y}) \in \operatorname{Lim}(g \circ f; \hat{x})$ do not imply the existence of $U_0 \in \rho(\hat{x})$ such that $g \mid f(U_0)$ is continuous at $\hat{y}$. This shows that our continuity assumption on g is a convenient sufficient but not a necessary condition for $g(\hat{y}) \in \operatorname{Lim}(g \circ f; \hat{x})$. Let be X, Y, Z the real line and, for every positive integer k, A_k the set of rational numbers in the interval $[2^{-k}, 2^{-k+1})$, $D_k = \{1/2^{k-1}(2l-1) \mid l$ a positive integer$\}$, and f_k any bijection from A_k onto D_k. We define $f: X \to Y$, $g: Y \to Z$ by

$$f(x) = \begin{cases} 0 & \text{if } x \in \mathbb{R} \setminus \bigcup_{k=1}^{\infty} A_k, \\ f_k(x) & \text{if } x \in A_k, \end{cases} \qquad g(y) = \begin{cases} 0 & \text{if } y \in \mathbb{R} \setminus \bigcup_{k=1}^{\infty} D_k, \\ 1/2^{k-1} & \text{if } y \in D_k. \end{cases}$$

Then we obtain for

$$U_k = (-1/2^{k-1}, 1/2^{k-1}) \setminus \{0\}\colon\ U_k \in \rho(0), f(U_k) = \{0\} \cup \bigcup_{j=k}^{\infty} D_j,$$

therefore $f(U_k) \subset [0, 1/2^{k-1}]$, i.e., $\operatorname{Lim}(f; 0) = \{0\}$ (uniqueness, cf. Remark 3). Since $y_l = 1/2^{k-1}(2l-1) \in D_k \subset f(U_k)$, $y_l \to 0$, $g(y_l) = 1/2^{k-1} \to 1/2^{k-1} (l \to \infty)$, $g(0) = 0$, $g \mid f(U_k)$ is not continuous at 0. So, for every $U \in \rho(0)$, $0 \in f(U)$ and $g \mid f(U)$ is not continuous at 0. But since $g[f(U_k)] \subset [0, 1/2^{k-1}]$, we have $\operatorname{Lim}(g \circ f; 0) = \{0\} = \{g(0)\}$.

REMARK 7. If X and Y are topological spaces and $f: X \to Y$ continuous and such that, for every $\hat{x} \in \hat{X}$, $\operatorname{Lim}(f; \hat{x})$ is a singleton, f need not possess a continuous extension to $\hat{X}$ (cf. [2]).

References

1. T. M. Apostol, Mathematical Analysis, 4th printing. Addison-Wesley, Reading, Mass., 1964.
2. N. Bourbaki and J. Dieudonné, Note de tératopologie II, Revue Scientifique (Revue Rose) 77 (1939) 180–181.
3. B. R. Gelbaum and J. M. H. Olmsted, Counterexamples in analysis, Holden-Day, San Francisco, 1964.
4. H.-J. Kowalsky, Topologische Räume. Mathematische Reihe, Band 26, Birkhäuser, Basel-Stuttgart, 1961.
5. E. J. McShane, A Theory of Limits. Studies in Modern Analysis, MAA Studies in Mathematics, Vol. 1 (1962) 7–29.
6. P. Ramankutty and M. K. Vamanamurthy, Limit of the composite of two functions, this MONTHLY, 82 (1975) 63–64.
7. W. J. Thron, Topological structures, Holt, Rinehart and Winston, New York, 1966.

MATHEMATISCHES INSTITUT DER UNIVERSITÄT BERN, SWITZERLAND.

MATHEMATICAL EDUCATION

EDITED BY PAUL T. MIELKE AND SHIRLEY HILL

Material for this Department should be sent to Paul T. Mielke, Department of Mathematics, Wabash College, Crawfordsville, IN 47933, or to Shirley Hill, Department of Mathematics, University of Missouri, Kansas City, MO 64110.

SOME TECHNIQUES FOR THE MATHEMATICS CLASSROOM

ALAN H. SCHOENFELD

This note summarizes a longer paper with the same title, available upon request from the author. Briefly, that article addresses the following problem: that there are numerous obstacles between the

dedicated teacher and his or her students in the classroom. The following is a list of some techniques my colleagues and I have found valuable for surmounting (or circumventing) those barriers.

A. Class evaluations. In the third or fourth week of the term you might give your students ten minutes to write a paragraph or two about their views of the class: pace and clarity of lectures, blackboard techniques, homework load, etc., are all fair game. At the next class meeting you might respond to these, explaining procedures they've challenged and announcing changes (if any). At the very least, the students are convinced of your good intentions and you find out about dissatisfaction (if there is any) before it's too late.

B. Alternate class structure. If the material is appropriate or if you planned to do some review, you can hand out a list of problems, have the class break into small groups (say 4 or 5 students each) and have the groups work individually on the problems for half of the class hour while you circulate among the groups acting as a consultant. In the second half of the hour you can present brief solutions and tie the material together. This approach both prevents boredom (everybody participates) and stimulates participation.

C. Student-written class notes. At each class meeting one student is designated "official note-taker." That person's responsibility is to write up a complete set of class notes for that day, perhaps in consultation with you. The notes are then copied and distributed to the class. Freed from the responsibility of taking their own notes, all but one student can participate more easily in class; you obtain a complete set of notes and the opportunity to see a different side of the students.

D. In-class "experiments." Unlike the physical sciences, we do not have laboratories, but we can still perform "experiments" of our own. The simultaneous solution of the equations

$$\begin{aligned} y'' &= -g; \qquad & y' &= V_0 \sin\theta; \qquad & y(0) &= 0 \\ x'' &= \ \ 0; \qquad & x' &= V_0 \cos\theta; \qquad & x(0) &= 0 \end{aligned}$$

and the effects of varying θ on the solutions can be demonstrated by tossing a piece of chalk through the classroom; a book crashing to the floor contradicts Zeno's paradox; Möbius strips fascinate students; and so on.

E. Quizzes without trauma. You might give brief, unannounced quizzes, to be taken anonymously. Examining the papers tells you what is or isn't getting across, at no "cost" to the students.

F. The class dummy. If you want participation and aren't getting it, you might assign or ask for a volunteer "dummy of the day." That person's job is to ask all the questions the others are afraid or ashamed to ask. When the students see that you welcome questions and respond reasonably, they start asking their own; and the "dummy" can be retired.

G. "Programmed" spontaneity. If you're doing routine problems, you can have the class participate in their construction. In a calculus class, for example, I ask the students to "give me some numbers." After they've called out "seven," "twelve," "pi," and "six point three," I ask them to differentiate $(7x^{12} + \pi)^{6.3}$. The break in routine can be entertaining.

H. "Treats." We're mathematicians because we like mathematics; our favorite problems or games can entertain students as well as ourselves. There are scores of mathematical amusements; why not share them with students if you have the time? (A particular favorite of mine is described in the larger paper.)

SESAME GROUP IN SCIENCE AND MATHEMATICS EDUCATION, C/O PHYSICS DEPARTMENT, UNIVERSITY OF CALIFORNIA, BERKELEY, BERKELEY, CA 94720.

A MATHEMATICS MINOR FOR PROSPECTIVE ELEMENTARY SCHOOL TEACHERS THAT SEEMS TO BE SUCCESSFUL

CHARLES R. MCNERNEY

The University of Northern Colorado is a university of 12,000 students. There are presently 1574 elementary education majors enrolled in the university who are required to take six quarter hours of mathematics as part of their general education program. The content of these six hours includes sets, logic, numeration, number theory, a development of the real and complex numbers and metrication. No geometry is taught in these two 3-hour courses. In addition to the above-mentioned six hours of mathematics, each elementary education major is required to take an academic minor.

The minor program. Prior to Fall 1972, the Mathematics Department offered a mathematics minor program which was the same for both secondary and elementary majors. This minor consisted of 30 hours of mathematics which included 10 hours of calculus, 5 hours of modern algebra, and 5 hours of college geometry. It goes without saying that very, very few elementary education majors elected to take this minor.

In the Spring of 1973, a recommendation was made to the department that a more realistic minor program be adopted, one more in line with the mathematical needs of the elementary school teachers. The case was argued that if an elementary school teacher needed to know calculus, the emphasis should be on the ideas of the calculus and the place of the invention of calculus in our Western culture. The rigor and the breadth of a regular beginning calculus course was not consonant with the needs of an elementary school teacher. Similar arguments were presented regarding abstract algebra and geometry. The program which was adopted was a 27 quarter-hour minor with a 3-hour informal geometry course. The remaining 24 hours of electives were to be planned between the minor advisor and the advisee. These hours were to be planned according to student mathematical needs and departmental objectives.

During the first year of the program, 31 students elected to take the minor. The program has shown growth in subsequent years to a present high of 75 students. There are also 6 students who have decided to take a split major in elementary education and mathematics. This major is somewhat less difficult than the major required of secondary mathematics teachers, but still a rigorous program.

Objectives of the program. There are at least three major objectives which should be attained in such a program. An obvious objective is to produce a student who is competent mathematically. It is felt that the student should have competency in the fundamentals of arithmetic, algebra, geometry, and trigonometry. In addition to this, the student should not only appreciate the aesthetic value of mathematics but also the utilitarian nature of the subject.

The content objectives are achieved in a number of ways. Most students who elect the program have studied one year of algebra and one year of geometry in high school. Those students who have only these mathematical competencies are counseled into a quarter of algebra which is equivalent to a high school Algebra II course. This course is followed with a trigonometry course and then a college algebra course. The remaining hours in the minor are spent in elective courses described elsewhere in this paper.

Those students who have had two years of high school algebra, a year of geometry and trigonometry, are counseled into a wide variety of courses which includes 6 hours of very rudimentary linear algebra, calculus, and probability. The emphasis in these courses is on business applications of the content. Linear functions become revenue functions, systems of equations become the model for break-even analysis, inequalities become constraints for linear programming and matrices are manipulated to solve inventory and marketing problems.

A foundations of arithmetic course is taught where students study topics in mathematics such as Fibonacci numbers, googols and googolplexes, palindromic numbers, phyllotaxis, and other topics

which are relevant to the elementary school curriculum. The presentation of most of the topics takes place in a mathematics lab setting.

The informal geometry course is a transformation geometry course, taught by means of tracing paper, geoboards, mirrors and the basic Euclidean constructions. After an intuitive development of the isometries, the student learns to use basic Euclidean constructions to express each isometry. Units on similarity and non-Euclidean geometry conclude the course. Time is spent in the course discussing the relative merits of the various pedagogical approaches to geometry.

Other elective courses include a course in elementary statistics, history of mathematics, computers and their impact upon society, mathematics of finance, elementary functions, and a mathematics activities course.

A second objective of the minor program is to build a positive attitude towards mathematics. It is our sincere belief that negative attitudes are transmitted to elementary school children either verbally or nonverbally by teachers possessing these attitudes. Likewise, a healthy, enthusiastic attitude towards mathematics will be transmitted to elementary school children. We feel that our program dispels the fear of mathematics that many of these students have, and slowly, but surely, nurtures a positive attitude. Students learn that mathematics can be fun. Hard work can be tolerated and even enjoyed if the pay-off is worthwhile.

A third objective of the program is to build a sense of professionalism in the student. Students are continually reminded they are expected to be leaders in mathematics education in their schools upon entry into the profession. The journals *The Arithmetic Teacher* and *The Mathematics Teacher* are used as supplemental reading in several courses.

Spin-offs of the program. The spin-offs of the program are noteworthy. First and foremost, teachers are entering the profession with considerable competence in mathematics.

There is also the advantage of increased enrollments in mathematics courses. This increased enrollment is generated by both math minor students and students who are not in the mathematics minor, but are interested in taking elective courses in mathematics. Courses are now available because of the minor program which appeals to the nonspecialist. Consequently, students are taking mathematics courses to partially fulfill their general education requirements.

With regard to research, several dissertations have been generated through the elementary mathematics experience. The research done in these courses has been both developmental and experimental.

A rather desirable spin-off of this program is a departmental awareness of the K–6 mathematics curriculum. To become more effective teachers, several members of the department are making an effort to familiarize themselves with the K–6 mathematics curriculum. These efforts provide a common ground for discussion between mathematicians and mathematics educators.

Several mathematics minors take courses above and beyond the 27 hours of mathematics that are required in the program. It seems that success breeds success. A basic premise under which our department works is that the students who are in the elementary math minor are mathematically competent, but that somewhere in their precollegiate education, many of these students were "turned-off" in mathematics. Our job is to "turn them on" and to encourage them to use their mathematical ability.

Future programs. Presently, we have two new programs in elementary mathematics being planned. One program is a course of study for teachers of the middle school. A second program is a revised major in mathematics for prospective elementary school teachers. We have every reason to believe that each program, when initiated, will be successful. When this happens, the next logical extension will be a graduate program for elementary school mathematics teachers.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF NORTHERN COLORADO, GREELEY, CO 80639.

RECREATIONAL MATHEMATICS FOR TEACHERS

JOHN L. HUNSUCKER

Many articles have been written in this journal and others on the problem of a small college mathematician remaining "alive." Another group of teachers of mathematics has this same problem, namely the teachers in our public school systems. Many of our mathematics departments are responsible to some degree or another for a large portion of the mathematics taught to our education students and therefore some thought should be directed toward this problem.

It was to this end that a course in recreational mathematics was taught at the University of Georgia in the Spring of '75. Note that recreational mathematics and "problem solving" are two different concepts. Problem solving refers more to the different methods of solving problems than to the problems themselves. Our course might more properly have been called "Fooling Around with Numbers and Things Mathematical."

The course was designed primarily for teachers of mathematics at the secondary or lower levels. It was felt that recreational mathematics might not only help provide these teachers and potential teachers with a means of staying active in mathematics but also provide background information for their teaching and a source of information for their gifted students.

Approximately three-fifths of the class were in-service teachers while the rest were senior or graduate level math education students. They came from all different teaching levels, from elementary through high school.

The class met in two two-hour sessions each week for ten weeks. About one hour of each session was devoted to giving the students problems and discussing the solutions to previous problems. The students presented most of the solutions. The works of Martin Gardner, W. W. R. Ball, Sam Loyd, and Henry Dudeney as well as Kordemsky's *Moscow Puzzle Book* were among the problem sources. Thirty minutes of each session were devoted to an individual student presenting a topic from recreational mathematics that he thought was applicable for the level at which he taught or would teach. The students' topics ranged all the way from paper folding at the elementary school level to polyominoes at the high school level. The rest of the period was devoted to students presenting solutions to problems that they had solved in various math and math ed journals. Some of the journals used were the *Monthly*, *School Science and Mathematics*, *Journal of Recreational Mathematics*, *Math Student*, and *Mathematics Magazine*. The students received a C for the thirty minute topic that they presented and then had their grade raised a letter for each problem from a journal that they solved successfully.

Two major problems arose in the presentation of this material. Finding problems that were interesting, challenging, and within the ability levels of students with such diversified backgrounds and interests, was one of the main problems. The other was locating problems in journals that the students with poorer backgrounds could solve.

When the course was over most of the students agreed that the objectives described earlier had at least been partially met. Some of them had even found that doing mathematics could be fun or at least entertaining.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF GEORGIA, ATHENS, GA 30602.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before April 30, 1977.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

E 2629*. *Proposed by David P. Robbins, Phillips Exeter Academy, New Hampshire*

Two points are chosen at random (uniform distribution) in the box $|x| \leqq a$, $|y| \leqq b$, $|z| \leqq c$ of $\mathbf{R}^3$. What is the expected distance between them?

(The cases when $\mathbf{R}^3$ is replaced by $\mathbf{R}$ or $\mathbf{R}^2$ are well known.)

E 2630. *Proposed by Edward T. Ordman, University of Kentucky*

Suppose that a polyhedral model (made, say, of cardboard) is slit along certain edges and unfolded to lie flat in the plane. The cuts may not be made so as to disconnect the figure. Now suppose that the resulting plane figure is again folded up to make a polyhedron (folding is allowed only on the original lines). The new polyhedron is not necessarily congruent to the original one. Find some interesting examples.

E 2631. *Proposed by Barry Powell, Kirkland, Washington*

In 1910 Mirimanoff proved that if p is an odd prime and $3^p \not\equiv 3 \pmod{p^2}$ then the equation

$$x^p + y^p = z^p$$

has no solution in positive integers x, y, z not divisible by p.

Show that the above condition is satisfied by all primes p having the form $p = \frac{1}{2}(3^{2^k} + 1)$ or $p = \frac{1}{2}(3^q - 1)$ with q also an odd prime.

E 2632. *Proposed by Azriel Rosenfeld, College Park, Maryland*

Define the *discrepancy* $d(A, B)$ between two plane geometric figures to be the area of their symmetric difference. Let A be a circle of radius r. Determine the inradius of the regular n-gon B for which $d(A, B)$ is minimal.

E 2633. *Proposed by Benjamin G. Klein, Davidson College, North Carolina*

Two points x and y in $\mathbf{Z}^n$ are said to be *neighbors* if $y - x = \pm e_i$ for some $i = 1, \ldots, n$ ($e_1, \ldots, e_n$ is the canonical basis of $\mathbf{Z}^n$). A subset $S \subset \mathbf{Z}^n$ is said to be *permutable* if there is a bijection $T: S \to S$ such that for each $x \in S$, Tx and x are neighbors. Show that if a finite subset $S \subset \mathbf{Z}^n$ is permutable then $|S|$ is even.

(*) Find necessary and sufficient conditions for a subset $S \subset \mathbf{Z}^2$ to be permutable.

E 2634. *Proposed by Jack Garfunkel, Forest Hills High School, Flushing, N.Y.*

Let A_i ($i = 0, 1, 2 \pmod 3$) be the vertices of a triangle, Γ its inscribed circle with center O. Let B_i be the intersection of the segment A_iO with Γ and let C_i be the intersection of the line A_iO with the side $A_{i-1}A_{i+1}$.

Prove that

$$\Sigma A_iC_i \leqq 3\Sigma A_iB_i.$$

SOLUTIONS OF ELEMENTARY PROBLEMS

Product of a Number and its Reverse

E 1243 [1956, 724; 1957, 434]. *Proposed by M. A. Rashid and M. A. Uppal, Lahore, Pakistan*

Prove that the product of a number consisting of two digits and its reverse is never a square except when the two integers are equal.

Comment by Greg Fitzgibbon, Valley Stream, N.Y. In the solution II (by Alan Wayne) the following conjecture was raised: When an integer (say x) and its reversal are unequal, their product is never a square except when both are squares. The following examples show that this conjecture is false.

$$8712 \cdot 2178 = 4356^2, \qquad 98901 \cdot 10989 = 21978^2.$$

One can also take for x any one of the following integers

$$x = 88(10^k - 1), \qquad k \geqq 2;$$

or $x = 8712 \cdot y$ where y has only 0 and 1 as digits, y is equal to its reverse, and any two consecutive digits 1 in y are separated by at least three zeros. See also the note by C. A. Grimm and Daniel W. Ballew: *Reversible multiples* in J. Recreational Math., 8(1975–76), 89–91.

The Knight's Distance

E 2392 [1973, 74; 1976, 379]. *Proposed by David Singmaster, London, England*

On the $n \times n$ chessboard, for $n \geqq 4$, define the *knight's distance* $D(A, B)$ between two squares A and B to be the minimum number of knight's moves required to go from A to B. Define the *knight's diameter* $M(n)$ of the $n \times n$ board to be the maximum knight's distance between any two squares of the board.

(1) Is $M(n)$ monotonic?

(2) Does $M(n)$ always equal the knight's distance between opposite corners of the board?

(3) Prove or disprove: For $n \geqq 5$, $M(n) = [2n/3]$.

(4) Determine the knight's distance $D(O, P)$ from the origin to an arbitrary square $P = (a, b)$ on the infinite chessboard.

III. *Comments by Roger Weitzenkamp, Oak Park, Illinois.* We shall prove that $M(n)=\{2n/3\}$, as originally stated by the proposer. We assume familiarity with the solution and notation as published previously.

LEMMA 1. *$M(n)\geqq\{2n/3\}$ for $n\geqq 5$.*

Proof. Case 1: $n=3k$. Then

$$\begin{aligned} M(n)&\geqq f(n-1,n-1)=f(3k-1,3k-1)\\ &=f(3(k-1)+2,\ 3(k-1)+2)=(3k-1)-(k-1)+1-1\\ &=2k=\{2n/3\}. \end{aligned}$$

Case 2: $n=3k+1$. Then

$$M(n)\geqq f(n-1,n-2)=f(3k,3k-1)=(3k-1)-(k-1)+1-0=2k+1=\{2n/3\}.$$

Case 3: $n=3k+2$. Then

$$M(n)\geqq f(n-1,n-1)=f(3k+1,3k+1)=(3k+1)-k+1-0=2k+2=\{2n/3\}.$$

LEMMA 2. *$M(n+3)\leqq M(n)+2$ for $n\geqq 4$.*

Proof. Let $D'(A,B)$ be the knight's distance between two squares A and B on an $(n+3)\times(n+3)$ chessboard X.

Let A and B be any two squares of X and let Y be a sub-board of X of size $n\times n$ such that Y lies in a corner of X and covers A. If Y also covers B then

$$D'(A,B)\leqq D(A,B)\leqq M(n).$$

If B is not covered by Y then there is a square C on Y such that $D'(C,B)\leqq 2$. Hence

$$D'(A,B)\leqq D'(A,C)+D'(C,B)\leqq D(A,C)+2\leqq M(n)+2.$$

Proof of $M(n)=\{2n/3\}$ for $n\geqq 5$. This is easy to verify when $n=5,6,7$. Assume this holds for some $n\geqq 5$. Then by Lemmas 2 and 3

$$M(n+3)\leqq M(n)+2=\{2n/3\}+2=\{2(n+3)/3\}\leqq M(n+3)$$

which yields $M(n+3)=\{2(n+3)/3\}$. By induction we are done.

This result disposes of parts (1) and (3) of the problem.

Cases 1 and 3 in the proof of Lemma 1 show that if $n\geqq 5$ the answer to (2) is affirmative if and only if $n\not\equiv 1 \pmod 3$.

The formula $M(n)=\{2n/3\}$ was also proved by Paul Campbell and Daniel Cohen (independently).

When 2^m-2^n divides 3^m-3^n

E 2468* [1974, 405; 1976, 288]. *Proposed by Harry Ruderman, Hunter College Campus School*

Suppose that $m>n\geqq 0$ are integers such that 2^m-2^n divides 3^m-3^n. Show that 2^m-2^n divides x^m-x^n for all natural numbers x.

II. *Remarks by The Mod Set Stanford University and Carl Pomerance, University of Georgia (independently).* Let A (resp. B) be the set of ordered pairs (n,k) where $n\geqq 0$ and $k\geqq 1$ are integers such that $2^n(2^k-1)$ divides $x^n(x^k-1)$ for all integers $x\geqq 1$ (resp. for $x=3$). We denote by A' (resp. B') the second projection of A (resp. B).

It was shown by A. Schinzel (*On primitive prime factors of a^n-b^n*, Proc. Cambridge Phil. Soc. 58 (1962), 555–562) that if $k\neq 1,2,4,6,12$ then 2^k-1 has a prime factor p greater than or equal to $2k+1$.

It follows from Velez's theorem that $k \notin A'$. Hence $A' = \{1,2,4,6,12\}$ in view of the list given by Velez.

Let $(n,k) \in A$. Then 2^n divides $3^k - 1$ so that for $k = 1$ we have $n \leqq 1$; for $k = 2$, $n \leqq 3$; for $k = 4$, $n \leqq 4$; for $k = 6$, $n \leqq 3$ and for $k = 12$, $n \leqq 4$. This gives 20 possible pairs (n, k). It is easy to rule out the 6 extraneous pairs

$$(0,2), \quad (0,4), \quad (0,6), \quad (1,6), \quad (0,12), \quad (1,12),$$

by taking $x = 3$. Thus A consists of 13 pairs listed by Velez plus the pair $(0,1)$ which was omitted by him as trivial.

The problem E 2468* can be stated as follows: Show that $B = A$. So far we have determined completely A; it consists of 14 pairs. By a computer search (at Stanford) it was verified that no integer k from 13 to 1900, inclusive, belongs to B'.

Reducing modulo 2

E 2552 [1975, 851]. *Proposed by P. Castevens, Winston-Salem, North Carolina*

Let A be an $n \times n$ real matrix with zeros on the main diagonal and ± 1 off the diagonal. Show that A is nonsingular if n is even but that A may be singular if n is odd.

Solution by James G. Mauldon, Amherst College. If n is even then $A \cdot {}^tA \equiv I_n \pmod 2$ which implies that A is nonsingular. (${}^tA$ is the transpose of A.) If n is odd and A is chosen skew-symmetric then it is singular.

Also solved by 84 other contributors and the proposer.

Editor's Comment. Several solvers noted that these also hold for any matrix with even integers on the diagonal and odd integers elsewhere. (The above proof applies.)

Simson and Euler Lines

E 2553 [1975, 851]. *Proposed by V. B. Sarma, Mylapore, India*

Suppose that A, B, C, D are concyclic and that the Simson line of A with respect to triangle BCD is perpendicular to the Euler line of triangle BCD. Show that the Simson line of B will be perpendicular to the Euler line of triangle CDA. Is the above result true if we replace "perpendicular" by "parallel"?

Solution by James G. Mauldon, Amherst College. We may assume that A, B, C, D lie on the unit circle in the complex plane. Let a, b, c, d be the complex numbers which correspond to these points. The Euler line p of triangle BCD goes through the origin and the orthocenter $b + c + d$ of BCD. The feet of the perpendiculars from A to BC, CD, DB are

$$x = \tfrac{1}{2}(a + b + c - \bar{a}bc), \quad y = \tfrac{1}{2}(a + c + d - \bar{a}cd), \quad z = \tfrac{1}{2}(a + d + b - \bar{a}db),$$

respectively. These points lie on the Simson line q of A with respect to BCD. The directions of p and q are given by $b + c + d$ and $2(x - y) = \bar{a}(a - c)(b - d)$. The lines p and q will be parallel (perpendicular) if and only if the complex number

$$e = \frac{\bar{a}(a-c)(b-d)}{b+c+d}$$

is purely imaginary (real). The conditions $\bar{e} = -e$ and $\bar{e} = e$ can be written in the form

(1) $bc + cd + db = a(b + c + d)$,

(2) $ab + ac + ad + bc + bd + cd = 0$,

respectively. Since (2) is symmetric in a, b, c, d the first part of the problem is finished.

The asymmetry of (1) indicates that the answer to the second part is negative. To prove this it

suffices to take

$$a=(12+5i)/13, \quad b=1, \quad c=i, \quad d=(3-4i)/5.$$

Also solved by M. G. Greening (Australia), Mark Kleiman, O. P. Lossers (Netherlands), Stephanie Sloyan, Daniel Sokolowsky, Albert Walker (Canada), and the proposer.

The first part of this problem appeared as advanced problem 4695 [1957, 437].

Polynomial Function Restricted to Rationals

E 2554 [1975, 851]. *Proposed by F. David Hammer, Stockton State College, New Jersey*

Can a polynomial function with integer coefficients be 1-1 when restricted to the rationals, but not 1-1 on the reals?

Solution by Lorraine L. Foster, California State University at Northridge. We claim that $f(x)=x^n-2x$ (n odd and $\geqq 3$) has the required properties. It is not one-to-one on the reals because it has three real roots. Assume that $f(x)=f(y)$ for some rationals x, y and write $x=a/b$, $y=c/d$ where a, b, c, d are integers with $b>0$, $d>0$, $(a, b)=(c, d)=1$. Then from $f(x)=f(y)$ we obtain

(1) $$d^n a(a^{n-1}-2b^{n-1})=b^n c(c^{n-1}-2d^{n-1}).$$

Since $(a, b)=1$ and $(a^{n-1}-2b^{n-1}, b)=1$ we must have $b^n \mid d^n$. Similarly, $d^n \mid b^n$. Therefore $b=d$ and (1) can be rewritten as

(2) $$a^n-c^n=2b^{n-1}(a-c).$$

Suppose $x \neq y$, i.e. $a \neq c$. Then (2) implies that

(3) $$a^{n-1}+a^{n-2}c+\cdots+ac^{n-2}+c^{n-1}=2b^{n-1}.$$

Since n is odd it follows from (3) that a and c must be even. But then the left-hand side of (3) is divisible by 4 forcing b to be even. This is a contradiction since $(a, b)=1$. This proves that we must have $x=y$, i.e., f is one-to-one on the rationals.

Also solved by Anders Bager (Denmark), David Bloom, Albert Briggs, Jr., John Bryant & Robert Gilmer & Andrew Tristram, J. D. Buckholtz, Leo Comerford, Jr., G. A. Heuer & Karl Heuer, A. A. Jagers (Netherlands), H. Kestelman (England), Ivan Korec (Czechoslovakia), J. C. Lagarias, S. C. Locke (Canada), James Mauldon, Donald Morrison, Michael Norris & William Vélez, William Nuesslein, Anthony O'Farrell (Ireland), Problem Solving Group Bern (Switzerland), H. Reuvers (Netherlands), Temple University Problem Solving Group, University of San Francisco Problem Group, University of South Alabama Problem Group, and the proposer.

Editor's comment. Most solvers used the example x^3-2x. Other examples which were used are $2x^3-x$, x^3-6x, $2x^3-5x$, x^3+2x^2-2x+1, x^6-3x^3.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Hill Center, Rutgers University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate signed sheets and should be mailed before April 30, 1977.

6126. *Proposed by Harold Reiter, University of North Carolina, Charlotte*

A well-known theorem of dimension theory states: The countable union of closed sets each of dimension zero has dimension zero.

If X is a metric space and $(2^X, D)$ is the associated space of compact subsets, with the Hausdorff metric, the following proposition, if true, would generalize the above theorem. Let S be a zero-dimensional collection of compact zero-dimensional sets. Then $\cup\{C: C \in S\}$ is zero-dimensional. Prove or disprove the proposition.

6127. *Proposed by M. J. Pelling, University of Benin, Nigeria*

Sum the series $\Sigma_{n=2}^{\infty}\zeta(n)(a/b)^n$ where $0<a/b<1$ is rational. (The answer should not involve the gamma function.)

6128. *Proposed by Martin Schechter and Peter Borwein, University of British Columbia*

Let 2^{ω} be the set of all sequences with entries 0 or 1 and let N^{ω} be the set of all sequences with entries from the non-negative integers. Can one construct a bijection f from 2^{ω} onto N^{ω} with the property that for any sequence X in 2^{ω}, one can compute the first n entries of $f(X)$ given only the first m entries of X (where m may depend on X and n)?

6129. *Proposed by E. H. Kronheimer, Birkbeck College, University of London, England*

Let S be a simple closed curve in the plane. Prove that, unless S is a circle, it is always possible to find four points p, q, u, v on S and a point x inside S, such that u and v belong to distinct components of $S\backslash\{p,q\}$, and x is nearer to both p and q than it is to either u or v.

6130. *Proposed by Erwin Just, Bronx Community College, and Eugene Levine, Adelphi University*

Prove that there exists a partition of the rational points of the plane into an infinite number of everywhere dense subsets such that each straight line containing two rational points will have a nonempty intersection with each of the subsets.

6131. *Proposed by Lee A. Rubel, University of Illinois, Urbana*

Suppose $\phi \geqq 0$ is in $L^1(-\infty,\infty)$, ϕ vanishes outside of $[a,b]$, and ϕ is strictly decreasing on $[a,b]$. Prove that the span of the translates of ϕ is dense in $L^1(-\infty,\infty)$.

SOLUTIONS OF ADVANCED PROBLEMS

A Measure in $\mathbb{R}^n$

5999 [1974, 1034; 1976, 491]. *Proposed by R. D. McKelvey, Carnegie-Mellon University and University of Rochester*

Let μ be a finite measure on the σ-algebra of sets in $\mathbb{R}^n$ generated by the half-spaces defined by hyperplanes through the origin. I.e., for any $\alpha \in \mathbb{R}^n$, let $H_\alpha=\{x\in\mathbb{R}^n \mid x\cdot\alpha>0\}$ and let $H=\{H_\alpha \mid \alpha\in\mathbb{R}^n\}$. Further, let B be the σ-algebra generated by H, and $\mu: B\to\mathbb{R}$ be a finite positive measure on B. Prove or disprove the conjecture: If $\mu(A)=\mu(-A)$ for all $A\in H$, then $\mu(A)=\mu(-A)$ for all $A\in B$. (Here, of course, $-A=\{x\in\mathbb{R}^n \mid -x\in A\}$.)

II. *Solution by D. A. Overdijk, F. H. Simons and F. W. Steutel, Technological University Eindhoven, The Netherlands.* We shall show that the conjecture is true. Without loss of generality we may assume $\mu\{0\}=0$. Then it is easy to construct an orthonormal basis $e_1,\ldots,e_n$ for $\mathbb{R}^n$ such that $\mu\{x\in\mathbb{R}^n \mid (x,e_n)=0\}=0$.

For every Borel set $Y\subset\mathbb{R}^{n-1}$ we define

$$Y^+=\{(\lambda x_1,\ldots,\lambda x_{n-1},\lambda) \mid (x_1,\ldots,x_{n-1})\in Y,\ \lambda>0\},$$

$$Y^-=\{(\lambda x_1,\ldots,\lambda x_{n-1},\lambda) \mid (x_1,\ldots,x_{n-1})\in Y,\ \lambda<0\}.$$

Obviously, $Y^+\in B$, $Y^-\in B$, $Y^+=-Y^-$, and the sets Y^+, Y^- generate B restricted to $\mathbb{R}^n\backslash\{x\in\mathbb{R}^n \mid (x,e_n)=0\}$.

Therefore, if we define $\mu^+(Y)=\mu(Y^+)$, $\mu^-(Y)=\mu(Y^-)$, then μ^+ and μ^- are finite positive measures on the Borel sets of $\mathbb{R}^{n-1}$ and the conjecture is true if and only if $\mu^+=\mu^-$.

Choose $a \in \mathbb{R}^n$, $b \in \mathbb{R}^n$ such that

$$\mu\{x \in \mathbb{R}^n \mid (x, a) = 0\} = \mu\{x \in \mathbb{R}^n \mid (x, b) = 0\} = 0,$$

and consider

$$A_1 = \{x \in \mathbb{R}^n \mid (x, a) > 0,\ (x, b) > 0\}$$
$$A_2 = \{x \in \mathbb{R}^n \mid (x, a) > 0,\ (x, b) < 0\}$$
$$A_3 = \{x \in \mathbb{R}^n \mid (x, a) < 0,\ (x, b) > 0\}$$
$$A_4 = \{x \in \mathbb{R}^n \mid (x, a) < 0,\ (x, b) < 0\}.$$

Then $A_1 = -A_4$, $A_2 = -A_3$. Moreover, from

$$\mu\{(x, a) > 0\} = \mu(A_1) + \mu(A_2) = \mu(A_3) + \mu(A_4) = \mu\{(x, a) < 0\}$$
$$\mu\{(x, b) > 0\} = \mu(A_1) + \mu(A_3) = \mu(A_2) + \mu(A_4) = \mu\{(x, b) < 0\}$$

we conclude

$$(*) \qquad \mu(A_1) = \mu(A_4) \quad \text{and} \quad \mu(A_2) = \mu(A_3).$$

For $u \neq 0$, $u \in \mathbb{R}^{n-1}$ and $\beta \in \mathbb{R}$ define

$$K(u, \beta) = \{x \in \mathbb{R}^{n-1} \mid (x, u) = \beta\}.$$

Since μ^+ and μ^- are finite measures and for $\beta_1 \neq \beta_2$, $K(u, \beta_1) \| K(u, \beta_2)$, we have $\mu^+ K(u, \beta) = \mu^- K(u, \beta) = 0$ for all but at most countable many β. Then for all those β, we have by (*)

$$\mu^+\{x \in \mathbb{R}^{n-1} \mid (x, u) < \beta\} = \mu^-\{x \in \mathbb{R}^{n-1} \mid (x, u) < \beta\},$$

hence even for all β this relation holds.

As a measure on the Borel sets of $\mathbb{R}^{n-1}$ is determined by its values of the sets $\{(x, u) < \beta\}$ (see, e.g., Loève, *Probability Theory*, 2nd edition, p. 205) it follows that $\mu^+ = \mu^-$.

Editor's Note. The proposer has written to point out a gap in the earlier Solution I which shows D to be closed only under disjoint unions, whereas for D to be a σ-algebra it must be shown that D is closed under arbitrary countable unions, and this is the core of the difficulty. Moreover, D is not a σ-algebra if the set H of generators is finite.

Algebraic Inequalities for π

6019 [1975, 307]. *Proposed by R. E. Shafer, Berkeley, California*

K. Chandrasekharan, *Introduction to Analytic Number Theory*, Chap. VII, Section 3, proves that

$$\frac{2^{2n}}{2\sqrt{n}} < \binom{2n}{n} < \frac{2^{2n}}{\sqrt{2n}}$$

for all integers $n \geqq 2$. Prove for all positive numbers n that

$$\frac{2^{2n}}{\sqrt{\pi}(n^2 + n/2 + 1/8)^{1/4}} < \binom{2n}{n} < \frac{2^{2n}}{\sqrt{(n + 1/4)\pi}}.$$

Solution by Emil Grosswald, Temple University. Both bounds may be proved by the same method. It seems, however, that the computation can be reduced by proceeding slightly differently for each inequality. We use the following well-known bounds, which follow immediately from Stirling's formulae as given, e.g., in M. Abramowitz and I. A. Segun, *Handbook of Mathematical Functions*, Dover Publ., Inc., New York, 5th ed., under 6.1.37 and 6.1.41 on p. 257:

$$n^{n+1/2}e^{-n}\sqrt{2\pi}\left(1+\frac{1}{12n}\right)<n!=\Gamma(n+1)<n^{n+1/2}e^{-n}\sqrt{2\pi}\left(1+\frac{1}{12n}+\frac{1}{288n^2}\right); \tag{1}$$

$$(n+\tfrac{1}{2})\log n-n+\tfrac{1}{2}\log 2\pi+\frac{1}{12n}-\frac{1}{360n^3}<\log n!=\log\Gamma(n+1)$$

$$<(n+\tfrac{1}{2})\log n-n+\tfrac{1}{2}\log 2\pi+\frac{1}{12n}-\frac{1}{360n^3}+\frac{1}{1260n^5}. \tag{2}$$

The inequality for the upper bound may be written as

$$\frac{(2n)!}{n!\,n!}<\frac{2^{2n}}{\sqrt{\pi n}(1+1/4n)^{1/2}}$$

or, using (1), after some simplifications, as

$$\left(1+\frac{1}{24n}+\frac{1}{4\cdot 288\,n^2}\right)\left(1+\frac{1}{6n}+\frac{1}{144\,n^2}\right)^{-1}<\left(1+\frac{1}{4n}\right)^{-1/2}.$$

By direct computations it follows that the left hand side is less than

$$1-\frac{1}{8n}+\frac{17}{4\cdot 288\,n^2}+\frac{7}{16\cdot 144\,n^3}+\frac{5}{16\cdot 12^3n^4}+\frac{1}{12^5n^5}+\frac{1}{8\cdot 12^6n^6},$$

while the right hand side exceeds

$$1-\frac{1}{8n}+\frac{3}{8\cdot 16n^2}-\frac{5}{16\cdot 64\,n^3}.$$

The inequality to be proved holds, therefore, provided that

$$\frac{17}{8\cdot 288\,n^2}+\frac{7}{16\cdot 144\,n^3}+\frac{5}{16\cdot 12^3n^4}+\frac{1}{12^5n^5}+\frac{1}{8\cdot 12^6n^6}$$

$$<\frac{3}{8\cdot 16\,n^2}-\frac{5}{16\cdot 64n^3}.$$

This is equivalent to

$$\frac{37}{2^8 3^2}>\frac{73}{2^{10}3^2n}+\frac{5}{2^{10}3^3n^2}+\frac{1}{2^{10}3^5n^3}+\frac{1}{2^{15}3^6n^4}$$

that is, equivalent to

$$37>\frac{73}{4n}+\frac{5}{12n^2}+\frac{1}{108n^3}+\frac{1}{2^7 3^4n^4}$$

which holds for all $n\geqq 1$. This finishes the proof for the upper bound.

The lower bound inequality is equivalent to

$$2n\log 2-\tfrac{1}{2}\log(\pi n)-\tfrac{1}{4}\log\left(1+\frac{1}{2n}+\frac{1}{8n^2}\right)<\log((2n)!)-2\log(n!).$$

By use of (2) this reduces to

$$-\frac{1}{8n}+\frac{1}{192\,n^3}-\frac{1}{630\,n^5}>-\frac{1}{4}\log\left(1+\frac{1}{2n}+\frac{1}{8n^2}\right),$$

which gives

$$\log\left(1+\frac{1}{2n}+\frac{1}{8n^2}\right)>\frac{1}{2n}-\frac{1}{48n^3}+\frac{2}{315n^5}.$$

The left hand side exceeds

$$\frac{1}{2n}-\frac{1}{48n^3}+\frac{1}{128n^4}-\frac{1}{128n^5}-\frac{1}{192n^6}-\frac{1}{1024n^7}-\frac{1}{2^{14}n^8}$$

so that it is sufficient to show that

$$\frac{1}{128n^4}-\frac{1}{128n^5}-\frac{1}{192n^6}-\frac{1}{1024n^7}-\frac{1}{2^{14}n^8}>\frac{2}{315n^5},$$

which reduces easily to

$$1>\frac{571}{315n}+\frac{2}{3n^2}+\frac{1}{8n^3}+\frac{1}{128n^4}.$$

This last inequality is evidently true for $n \geqq 3$; for $n = 1$ and $n = 2$ we verify the proposed lower bound directly: Thus the lower bound is established.

It may be worthwhile to observe that the lower bound is very accurate; already when $n = 6$, $\frac{1}{4}\binom{2n}{n} = 231$, while the corresponding bound is about 230.9997.

Also solved by M. T. Bird, Paul Bruckman, Joseph Grimland Jr. & Samuel Glidewell, O. P. Lossers (Netherlands), Ram Murty & Kumar Murty (Canada), F. G. Schmitt, Jr., Philip Young and the proposer.

Editor's notes. Grimland and Glidewell provide an elegant upper bound for $\binom{2n}{n}$ namely $2^{2n}\pi^{-1/2}(n^2+n/2+3/32)^{-1/4}$.

The proposer, in a manuscript using continued fractions, has extended his results and proved the inequality:

$$\frac{\Gamma(x+(1+h)/2)}{\Gamma(x+(1-h)/2)}<\left\{x^2+\frac{1-h^2}{12}\right\}^{h/2}, \qquad x>0,\ 0<h<1.$$

Friendly Integers

6020 [1975, 307]. *Proposed by C. W. Anderson and Dean Hickerson, University of California, Berkeley*

A pair of distinct natural numbers (k, m) is called a friendly pair (k is a friend of m) if $\Sigma(k)=\Sigma(m)$, where $\Sigma(n)=\sigma(n)/n$, where $\sigma(n)$ is the sum of the divisors of n. (For example, (4320, 4680) is a friendly pair.) Show that almost all numbers have friends, i.e., the natural (asymptotic) density of numbers with friends is unity. Equivalently, the density of solitary numbers (numbers without friends) is zero.

I. *Comment by M. G. Greening, University of New South Wales, Australia.* We prove that numbers n for which $(n, \sigma(n))=1$ are solitary.

Proof. (i) $\Sigma(p^\alpha)$ is an increasing function of d as $\Sigma(p^{\alpha+1})-\Sigma(p^\alpha)=p^{-(\alpha+1)}$,

(ii) $\Sigma(n)>1$, all $n>1$.

Consequently if $(n, \sigma(n))=1$, then $\Sigma(n)$ is a reduced fraction and any friend m must be a multiple of n. But if $m = hkn$ with $(h, n)=1$ and $(k, n)>1$, we have $\Sigma(m)>\Sigma(kn)$ by (ii), unless $h = 1$.

If $m = kn$, then $\Sigma(m)=\Sigma_{i=1}^r(p_i^{\beta_i})$ where $n=\Pi_{i=1}^r p_i^{\alpha_i}$ and $\beta_i \geqq \alpha_i$. So

$$\Sigma(m)=\prod_{i=1}^r \Sigma(p_i^{\beta_i})>\prod_{i=1}^r \Sigma(p_i^{\alpha_i})$$

by (i), unless $k = 1$. Thus the set of solitary numbers includes all primes and powers of primes.

II. *Comments by the proposers.* (1) Pairs of even friendly numbers are easy to find. Neville Robbins

found the pair (135, 819) of odd friendly numbers. Another pair may be found by starting with the quadruply perfect numbers $a = 2^5 \cdot 3^3 \cdot 5 \cdot 7$ and $b = 2^5 \cdot 3^4 \cdot 7^2 \cdot 11^2 \cdot 19^2 \cdot 127$ discovered by Descartes (1638) and Carmichael and Mason (1911), respectively. Since they are divisible by the same powers of 2, we have $945 = 3^3 \cdot 5 \cdot 7$ and $3^4 \cdot 7^2 \cdot 11^2 \cdot 19^2 \cdot 127$ are a friendly pair.

(2) For a friendly pair of opposite parity we have 42 and $3^2 \cdot 5 \cdot 7^2 \cdot 13 \cdot 19$. Also

$$\Sigma(2 \cdot 3^6 \cdot 23 \cdot 137 \cdot 547 \cdot 1093) = \Sigma(3^4 \cdot 5 \cdot 7 \cdot 11^2 \cdot 19)$$

(3) Known solitary numbers not greater than 100 are

1	2	3	4	5	7	8	9
11	13	16	17	18	19	21	23
25	27	29	31	32	35	36	37
41	43	47	48	49	50	52	53
55	57	59	61	63	64	65	67
71	73	75	77	79	81	83	85
89	93	97	98	100			

There are 53 of these.

Numbers not greater than 100 and known to have friends are:

6	12	24	28	30	40	42	60	66	80	84

There are 11 of these. The situation for the remaining 36 numbers is presently unknown.

From this meager numerical evidence, it may seem that the solitary numbers are numerous in **N**; but the density of numbers n such that $\mathrm{GCD}(n, \sigma(n)) = 1$ equals zero.

Some solitary numbers n satisfy $\mathrm{GCD}(n, \sigma(n)) > 1$. 18, 48, and 52 are such. Consider 18: $\mathrm{GCD}(18, \sigma(18)) = 3$ and $\Sigma(18) = 13/6$. Any solution of $13n = 6\sigma(n)$ satisfies $n = 2^a 3^b m$ with $\mathrm{GCD}(6, m) = 1$. If $a \geqq 2$, then $\Sigma(n) \geqq \Sigma(2^2 3) = 7/3 > 13/6$. Hence, $a = 1$. If $b = 1$, then $\Sigma(m) = 13/12$, which has no solutions. $a = 1$, $b = 2$, $m = 1$ (i.e., $n = 18$) satisfies $\Sigma(n) = 13/6$. Moreover, any multiple of 18 except 18 is mapped under Σ to a rational $a/b > 13/6$. Therefore, $\Sigma(n) = 13/6$ has the single solution $n = 18$. Similar considerations show that 48 and 52 are solitary.

The density of friendly numbers is positive. Where $\mathrm{GCD}(n, 42) = 1$, we have $\Sigma(6n) = \Sigma(28n) = 2\Sigma(n)$. The density of n coprime with 42 is 2/7. Where $kA = \{ka \mid a \in A\}$ and when $d(A)$ the density of A exists, then $d(kA) = d(A)/k$. Hence the density of friendly numbers is no less than

$$\frac{8}{147} = \frac{2}{7}\left[\frac{1}{6} + \frac{1}{28} - \frac{1}{84}\right].$$

Linear Dependence in l^p

6021 [1975, 307]. *Proposed by C. R. Diminnie and Albert White, St. Bonaventure University*

In l^p, $p > 2$, does $\|x - y\| \|x + y\| = |\|x\|^2 - \|y\|^2|$, with $x, y \neq 0$, imply that $y = \alpha x$ for some real α?

Solution by John Lagnese, Georgetown University. The result is not true in l^1 or l^∞ but is true in l^p for $1 < p < \infty$. In fact, one can prove: *Let X be a strictly convex Banach space. If x and y are nonzero elements of X which satisfy*

$$\|x - y\| \|x + y\| = |\|x\|^2 - \|y\|^2|,$$

then x and y are linearly dependent.

A strictly convex space is characterized by the inequality $\| u+v \| < \| u \| + \| v \|$ for any two linearly independent elements u, v of X. It follows from the proof of the Minkowski inequality that the l^p spaces $(1<p<\infty)$ are strictly convex. Any reflexive Banach space can be made strictly convex by renorming the space with an appropriate, equivalent norm.

One may assume that $\| x \| \geqq \| y \|$ and that $y \neq \pm x$. Introducing $z = x + y$, $w = x - y$, the result to be proved may then be stated: If z and w are nonzero elements of X which satisfy

$$\| z+w \|^2 - 4\| z \| \, \| w \| - \| z-w \|^2 = 0,$$

then z and w are linearly dependent. Otherwise, one obtains

$$4\| z \| \, \| w \| + \| z-w \|^2 = \| z+w \|^2 < (\| z \| + \| w \|)^2,$$

that is $\| z-w \| < |\| z \| - \| w \||$, which contradicts the triangle inequality.

That the result is not true in l^1 or l^∞ can be seen by choosing $x = (\frac{1}{2}, 1, 0, \cdots)$ and $y = (0, \frac{1}{2}, 0, \cdots)$.

Also solved by John Connett, C. Martin & J. E. Valentine, Bruce Reznick, D. A. Robbins, and John Swetits.

Minimal Intersection in a Collection of Sets

6022. [1975, 308]. *Proposed by Neal Felsinger, Yale University*

Given a collection X of subsets of S, no one containing another, let $C(X)$ consist of all minimal subsets of S which intersect every member of X. (For properties of $C(X)$ see Problem 5883 [1974, 293].) Show that if S is infinite, $C(X)$ does not necessarily exist.

Solution by James Baumgartner, Dartmouth College. Let N be the set of positive integers, and let $S = N \times N$. For each $n \in N$ let $A_n = \{(p,q) : p \neq n \text{ and } q \geqq n\}$. Let $X = \{A_n : n \in N\}$. Note that if $m \neq n$ then $A_m \not\subseteq A_n$. Suppose B intersects every A_n. Then B must be infinite. If there exists n_0 such that $\{q : (n_0, q) \in B\}$ is infinite, then let $(n_0, q) \in B$ and let $B' = B - \{(n_0, q)\}$. If no such n_0 exists, let $(p,q) \in B$ be arbitrary and let $B' = B - \{(p,q)\}$. In either case, B' still intersects every A_n. This shows that $C(X)$ must be empty.

Also solved by R. Duke & A. Vella (England), M. J. Pelling (Nigeria), Ellen Hertz, Peter Slater, and Richard Stanley & the MIT Combinatorial Seminar, 1974.

Borel Sets in a Product Space

6023 [1975, 308]. *Proposed by S. J. Sidney, University of Connecticut*

If for each k in the uncountable index set K, I_k denotes a copy of $[0,1]$ and U_k denotes the copy of $(0,1]$ contained therein, prove or disprove that $\Pi_k U_k$ is a Borel set in the compact space $\Pi_k I_k$.

Solution by D. Ž. Djoković, University of Waterloo, Canada. For each subset $S \subset K$ let $I(S) = \Pi_{k \in S} I_k$ and let $A(S)$ be the set of all subsets of $I(K)$ of the form $P \times I(K-S)$ where $P \subset I(S)$ and we identify canonically $I(S) \times I(K-S)$ with $I(K)$. Let $\mathscr{A}$ be the union of all $A(S)$ for countable subsets S of K. It follows that $\mathscr{A}$ is a σ-algebra containing the Borel sets of $I(K)$ and that $\Pi_k U_k \notin \mathscr{A}$.

Also solved by James Baumgartner, G. A. Edgar, M. J. Pelling (Nigeria), and the proposer.

Notes. (1) Baumgartner and Edgar show that $U = \Pi_k U_k$ is not almost open (property of Baire). (2) Pelling and the proposer solve the problem by showing that U is not measurable, after first constructing a suitable measure.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN
with the assistance of the mathematics departments of St. Olaf and Carleton Colleges
COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, CARLETON COLLEGE

We invite readers to submit reviews of significant recent college-level mathematics books. We especially encourage reviews based on classroom use, or comparative reviews of several related books. Reviews should ordinarily not exceed two pages (per book) typed double spaced. Manuscripts of reviews as well as books submitted for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057.

Methods in Classical and Functional Analysis. By Einar Hille. Addison-Wesley, Reading, Massachusetts, 1972. ix + 486 pp. $16.95. (Telegraphic Review, May 1973.)
Functional Analysis in Historical Perspective. By A. F. Monna. Halsted Press, Wiley, New York, 1973. viii + 167 pp. $16.50. (Telegraphic Review, May 1974.)
Principles of Functional Analysis. By Martin Schechter. Academic Press, New York, 1971. xix + 383 pp. $20. Student Edition, 1975, $16. (Telegraphic Review, January 1973.)

The books by Hille and Schechter are textbooks *par excellence*; Monna's seems to be the first book which is completely devoted to the history of functional analysis. The three books differ in scope and in many technical details. However, there are several philosophical threads that pull them together: a common outlook on the role and purpose of functional analysis; a return to the roots of the subject; and a stress on mathematical exposition.

We quote from the preface to Hille's book, "Modes come and go in mathematics as in most fields. During the half-century and more that I have worked in the vineyard I have heard many dire predictions for the fate of my ideas and interests. Abstraction has been in the saddle during most of the time and has ridden us mercilessly. In a modest way I have taken part in this development. I did not believe in abstraction *per se*; one should know what one is trying to generalize and one should show that the generalization is significant. I have tried to keep at least one foot on the ground while craning my neck to look into Heaven. Was it Heaven? There are some doubts, and the more extravagant claims of abstract mathematicians to be the sole dispensers of the true faith and the arbiters of values are received with a healthy skepticism." Hille says that his book may be regarded as "part of the backlash." The book certainly has a thesis; it is that a functional analyst is an analyst, first and foremost, and not a degenerate species of topologist. His problems come from analysis and his results should throw light on analysis.

Hille's thesis needs no defense; it suffices to look at functional analysis in historical perspective. This is where Monna's book comes in handy. It provides a readable survey of early development of functional analysis, and sketches significant features of this development by treating some characteristic examples. Concepts such as Hilbert space, Banach space, weak topology, completely continuous operators, etc., are nowadays common tools. How did they enter mathematics? Origins lie in the 19th century — in problems of mechanics, physics and classical analysis. It is refreshing to be reminded that functional analysis did not develop simply as a result of the tendency towards axiomatization, but that the need for abstractness evolved from subjects which gave rise to the development of functional analysis. These sources include the theory of infinite systems of linear equations, integral equations, calculus of variations, theory of best approximations, moment problems, quantum mechanics, generalized functions and operational calculus. Chapter I of Monna's book treats the role of some of these sources in the development of functional analysis.

Seen against the background of analysis as it was eighty years ago, three characteristics are dominant in functional analysis. They can briefly be described as (1) a tendency towards algebraiza-

tion in analysis (witness the modern version of the Weierstrass–Stone theorem, function algebras and ideals, etc.); (2) a stream towards results of structural nature; (3) the strong influence of topology (the origin of several kinds of topologies, e.g., the weak and weak-star topologies, resides in function spaces of analysis). Monna traces in Chapters II and III the introduction and great tracks of developments of linear spaces and normed spaces in analysis, and distinguishes two lines of approach. The first line leads from Fredholm via Hilbert, Schmidt, Riesz, Helly, to Hahn and Banach and the Polish school when explosive developments of functional analysis started. These developments (which led to the axiomatic theory of normed spaces) lie mainly in the period 1900 to 1930, culminating in 1932 with the publication of Banach's famous book, *Théorie des opérations linéaires.* The second line starts from the Italian mathematicians Peano, Pincherle and Volterra, who refer in their work to Laguerre and Grassmann. Their approach was established for the most part before 1900 and it has two culminating points: (i) the introduction of the concept of an abstract vector space by Peano (1888) and by Pincherle ($\pm$1890), (ii) the emergence of the French school of Hadamard and Fréchet whose work — as far as functionals and general analysis are concerned — is based on that of the Italian school. The contributions of Fréchet and E. H. Moore to general analysis are of particular significance.

The historical perspective in Monna's book is not carried beyond 1932. It stresses abstract spaces and does not deal much with other aspects and theorems of functional analysis, with the exception of the Hahn-Banach theorem, whose history is adequately traced in the book.

The many quotations throughout the book (nearly one third of the book) from papers by the mathematicians cited earlier and by others, add considerably to an appreciation of the perspectives. Profound knowledge of the subject is not required; any student with a certain amount of mathematical maturity will be able to read Monna's book. As such the book can be used for historical background of modern mathematics as well as for collateral reading in functional analysis courses. Unfortunately the book is much overpriced (84 cents/ft^2 of paper). Could it be that the lessons of history are always expensive?

Hille's book is suitable as an advanced undergraduate and graduate-level text. It is elementary in the sense that the prerequisites are confined to an exposure to undergraduate analysis; linear algebra is developed in Chapter 1. Yet the book leads by simple generalizations to advanced mathematics, and treats topics not usually covered in similar books at this level. The author emphasizes methods of construction rather than general theory and frequently sacrifices generality for simplicity in statements and proofs. General ideas are applied to illustrative material taken from basic infinite-dimensional spaces such as continuous functions, functions of bounded variation, sequences, and Lebesgue spaces. Applications of functional analysis to analysis are stressed and unvarnished classical analysis is handed out in significant doses. The book gives the reader a historical perspective and shows how the multitude of abstract concepts have arisen and are present *in nuce* already in Euclidean spaces. Historical remarks are dispersed throughout the book including names and life intervals of mathematicians who brought the ideas of the specific subject into life.

Three varieties of topics are treated in Hille's book. First, there are the usual canonical items: linear transformations, inner product spaces, the spectral theorem for Hermitian operators and operational calculus, Banach spaces and algebras, the weak topology, reflexive spaces, and some of the standard theorems of functional analysis including the Riesz representation theorem, the principle of uniform boundedness, the Hahn-Banach theorem, the Banach-Steinhaus theorem, and the closed graph theorem. Secondly, there are topics that are not always treated in elementary functional analysis books: Fourier series; several variants of the contraction mapping theorem and the implicit function theorem with many and novel applications; complex analysis in Banach spaces and Banach algebras, where emphasis is placed on the close relationship between classical theory of functions of complex variables and abstract theory of such functions. Finally there are many novelties and fascinating topics that do not seem to have been treated in similar textbooks. E. Landau's inequality (1913) for real functions $\|f'\|^2 \leqq 4\|f\|\,\|f''\|$ appears as a special case of a fundamental result of R. Kallman and G. C. Rota (1967): $\|Ax\|^2 \leqq 4\|x\|\,\|A^2x\|$, where A is the infinitesimal generator of any

contraction semigroup. (The paper by I. J. Schoenberg on the elementary cases of Landau's problem, this MONTHLY 80 (1973), 121–158 is highly recommended in this connection.) Two chapters are devoted to functional inequalities including celebrated results of Carathéodory, Nagumo, and Gronwall's inequality, with applications to differential equations and novel uses of convex functions and fixed point methods. The chapter on functional equations is particularly fascinating. Hamel basis is introduced where it first cropped up: the Cauchy functional equation. Many other ideas are traced to their roots. This chapter has a particular theme: "Mathematics is full of codes which have to be deciphered. Thus any functional equation contains a message more or less well hidden. All properties of the solutions are built into the equation and our problem is to bring them out by asking appropriate questions." An effort is made to show what type of *a priori* information is obtainable and how to obtain it. This technique is called "cryptoanalysis," an example of the picturesque language of the author. The last chapter treats a class of mean values (an abstract notion which includes as special cases the arithmetic, geometric, and the power means) and provides connections with functional equations and inequalities. Applications are given to geometric and extremal problems, to two classes of set functions: transfinite diameters and Čebyšev constants, and to potential theory.

Throughout the text, worked-out examples are given. Each section has exercises, with problems in the final sections designed to inspire the student to further study. In all there are over 850 exercises, a fair part of which are byproducts of the author's own research.

Do you remember Goursat's *Cours d'Analyse Mathématique* or de la Vallée-Poussin's *Cours d'Analyse Infinitésimale*? (These two classics left lasting impressions on generations of students. The English translation of Goursat's "Course" was undertaken at the suggestion of W. F. Osgood, whose lengthy review of the French edition appeared in *Bull. Amer. Math. Soc.* 9 (1903), pp. 547–555, under the headline "A Modern French Calculus," and cited the lack of standard texts on mathematical subjects in the English language.) Things have changed. Einar Hille has produced a new gem that will remain on the shelf of classics in analysis for many years and will serve as a model for mathematical exposition, the ingredients of which are described by P. R. Halmos: "Machines can write theorems and proofs, and read them. The purpose of mathematical exposition is for people to communicate ideas, not theorems and proofs. Experience shows that almost always the best way to communicate a mathematical idea is to talk about concrete examples and unsolved problems." Hille claims that his book is not polished; there are loose ends and many unsolved problems waiting for the craftsman. Instead of disputing this claim, let us wish there were more such unpolished books.

Hille's book is a return to moderation rather than a lavish backlash. It is not a resurrection of Whittaker and Watson's *Modern Analysis.* It is *not* two books, one on classical analysis and another on functional analysis, bound together under one cover at the suggestion of a publisher.

Where does Hille's book fit into the curriculum? The material in the book can serve different purposes. Chapters 1 to 4 and 7 to 11 can serve as a text for an introductory course in functional analysis. Chapters 5, 6, and 12 through 15 stand closer to classical analysis than does the rest of the book and can serve as a text for a course in aspects of analysis. However, this reviewer recommends the use of the entire book as a basic year-course in analysis. A modern *Cours d'Analyse Mathématique* should give the student a tour of the global paths and major landmarks in both functional and classical analysis, and present a unified picture with a few significant applications. An eminent mathematician and a masterful expositor has now made such a tour possible and enjoyable.

Yet Hille's book does not seem to be widely used at present as a textbook. The book does not purport to follow any particular CUPM recommendations; it does not easily fit into the current compartmentization of courses in our curriculum. But perhaps we have been too dogmatic as to what constitutes a proper mathematics curriculum. One way to design a curriculum is by identifying a set of textbooks. Hille's book could define a "core" of analysis in a mathematics curriculum.

We turn next to Schechter's book. Einar Hille reviewed it in the *American Scientist* and said: "'Charming' is a word that seldom comes to the mind of a science reviewer but if he is charmed by a

treatise, why not say so? I am charmed by this book." This reviewer used Schechter's book in a one-year course in functional analysis for beginning graduate students and senior undergraduates. The entire book was covered; the last five weeks were devoted to selected topics in nonlinear functional analysis from a book edited by L. B. Rall. On subsequent occasions, Schechter's book was also used in a 5-hour quarter course, covering several chapters. In both cases the book was found eminently suitable as a student-oriented textbook. Much of the book can be understood by a student having taken a course in advanced calculus. However, in several chapters an elementary knowledge of functions of a complex variable is required. Only rudimentary topological or algebraic concepts are used; they are introduced as needed. No measure theory is employed or mentioned. Attention is restricted to normed vector spaces and their important examples: Banach and Hilbert spaces. The book incorporates new material and includes some topics which are not usually found in textbooks on functional analysis, such as a comprehensive treatment of Fredholm and unbounded semi-Fredholm operators, essential spectrum, hypernormal and semi-normal operators, numerical range of an unbounded operator, and specific types of differential operators on $L^2(-\infty,\infty)$.

There are fourteen chapters: Basic Notions; Duality; Linear Operators; The Riesz Theory of Compact Operators; Fredholm Operators; Spectral Theory; Unbounded Operators; Reflexive Banach Spaces; Banach Algebras; Semigroups; Hilbert Space; Bilinear Forms; Self-adjoint Operators; Examples and Applications. Each chapter has a set of problems. On the other hand, the book does not introduce the weak topology, although weak convergence is briefly discussed. The author places emphasis on insight and heuristic ideas before he proceeds to formalism and generalizations.

Those who view functional analysis as a fertile tree of topological spaces with Banach and Hilbert spaces lying somewhere under the branch of barreled spaces most likely would not (and should not) choose Schechter's book as a text. Those who feel that the primary emphasis in a first course in functional analysis should be on operators in the context of Banach and Hilbert spaces and on problems of analysis, should find this book charming. Schechter's book is probably more suitable than Hille's book for such a course, while the latter is an excellent choice for a course with a broader view of analysis.

The writing of a book is often a labor of love and a reflection of an experience. Yet book reviews usually deal only with the finished product — the book itself (as if they are caving in to the domination of bureaucratic impersonality). The experience of the authors of the books under review had a great deal to do with the finished products. Hille walked the road of classical and functional analysis and left significant imprints; his book has been lecture material given from Bombay to Kingston, R. I. with a few stops on the way (in addition to the long tenure at New Haven), before the final book was written at Albuquerque. Schechter worked on problems of spectra of partial differential operators, Fredholm operators, and related matters; his choice of topics is influenced by strong interaction with concrete analysis. Monna worked in several areas of functional analysis (e.g., his book *Analyse Non-Archimédienne*) and is a prolific contributor to historical aspects of modern mathematics. Obviously they all love writing.

M. Z. NASHED, Georgia Institute of Technology

TELEGRAPHIC REVIEWS

Telegraphic reviews are designed to give prompt notice of new books with sufficient information to assist our readers in deciding whether to order an examination copy or to suggest library purchase. Possible uses are indicated as follows:

T = textbook P = professional reading
S = supplementary reading L = undergraduate library purchase
13 to 18 = freshman to second year graduate level usage
1 to 4 = appropriate time in semesters to cover text

Asterisks (*) or question marks (?) denote special positive or negative emphasis, respectively. Publishers are denoted by standard abbreviations; complete addresses may be found in *Books in Print*.

GENERAL, T(13-18), S*, P*, L**. *Companion to Concrete Mathematics, V. II: Mathematical Ideas, Modeling and Applications*. Z.A. Melzak. Wiley, 1976, xvi + 413 pp, $29.95. More examples drawn from a wide range of human activity serving to illustrate the use of mathematics and mathematical thinking in solving problems. A continuation of the high quality established in the first volume (TR, March 1974). LCL

GENERAL, S*(13), L*. *New Recreations with Magic Squares*. William H. Benson, Oswald Jacoby. Dover, 1976, viii + 198 pp, $4 (P). Comprehensive: includes classification, presentation of methods (many new--e.g., a cyclical method applicable to both even and odd squares), problems of enumeration, and final chapter giving proofs (good examples of modular arithmetic). Presumes only high school algebra; reads like Ball and Coxeter's *Mathematical Recreations*. LCL

GENERAL, T(13-18: 1, 2), L. *Patterns of Problem Solving*. Moshe F. Rubinstein. P-H, 1975, xvi + 544 pp, $15.95. Developed from a campus-wide interdisciplinary course. Conceives of problem-solving as a dynamic process involving a variety of disciplines. Treats language, computers, probability, information, modelling, cybernetics, optimization, without much motivation of their connection with problem-solving or with each other. Notable emphasis on attitudes and sharp focus on values in problem-solving. Instructors' manual and solutions manual available. PJC

BASIC, T(13: 1), S. *Beginning Algebra for College Students*. Karl J. Smith, Patrick J. Boyle. Brooks/Cole, 1976, xi + 395 pp, $11.95. Covers the first two years of high school algebra, no previous knowledge of which is assumed. Numerous, very helpful, attractively presented graphs, illustrations, and cartoons. Explanations clear and arranged well on the page. An enjoyable book to work through. Exercises. Appendices. Answers to selected problems. Index. RJA

PRECALCULUS, T(13: 1). *Elementary Functions, Second Edition*. H.S. Bear. Page-Ficklin, 1976, 413 pp, $7.50 (P). Informal treatment of topics from analytic geometry, transcendental functions, and vector analysis. Slight revisions from first edition (TR, April 1972), mostly in examples and addition of answers to review questions. LLK

EDUCATION, T(13-15: 1, 2), S. *Elementary Mathematics*. Donald F. Devine, Jerome E. Kaufmann. Wiley, 1977, xiii + 525 pp, $13.95. For prospective teachers of elementary mathematics. Sets, numeration, numbers (integers, rationals and irrationals), plane geometry (including area and volume), probability and statistics. Concrete, informal, intuitive appeal; conversational. Problem sets include discussion questions (concerning teaching methodology) and problems for the calculator. LCL

HISTORY, P. *Oeuvres de Paul Painlevé, Tome III*. Paul Painlevé. Ed Cent Nat Rech Scient, 1975, 826 pp, 180 F. Paul Painlevé was a mathematician, physicist, politician, and aviator (among other things). The tome contains his papers on second order differential equations, mechanics, letters, remembrances of Painlevé and a picture of Painlevé welcoming Lindbergh after the flight. SG

HISTORY, T(13-16: 1), S(10-16). *The Historical Roots of Elementary Mathematics*. Lucas N.H. Bunt, Phillip S. Jones, Jack D. Bedient. P-H, 1976, xii + 299 pp, $11.95. A pleasant ramble, which will delight a broad spectrum of general readers. Five chapters treat Greek mathematics with one each to Egyptian, Babylonian, and post-Greek. Incorporates material from Bunt's *Fan Ahmes tot Euclides* and Jones' *Understanding Numbers: Their History and Use*. (The authors fail to mention that the Sylvester process (p. 17) had been described earlier by Fibonacci.) PJC

FOUNDATIONS, S*(15-18), P**, L**. *Proofs and Refutations: The Logic of Mathematical Discovery*. Imre Lakatos. Cambridge U Pr, 1976, xii + 174 pp, $4.95 (P); $19.50. Modified and extended version of an essay in *British Journal of the Philosophy of Science*, 14 (1963/64). Examines the methodology of mathematics by means of a sharp dialectical dialogue concerning the Descartes-Euler conjecture (V-E+F=2 for polyhedra)--the history of which is exposed in great detail--plus further remarks on 19th century analysis. The author opposes formalism in mathematical writing and deductivist style in teaching, asserting a heuristic pattern for growth of informal mathematical theories. "Its modest aim is to elaborate the point that informal, quasi-empirical, mathematics does not grow through a monotonous increase of the number of indubitably established theorems but through the incessant improvement of guesses by speculation and criticism, by the logic of proofs and refutations." A splendid companion to Pólya's books on heuristics. PJC

FOUNDATIONS, T(15-17), S*, L. *Exercises in Set Theory*. L.E. Sigler. Springer-Verlag, 1976, 134 pp, $3.95 (P). Useful collection of exercises (with solutions) exploring definitions and theorems of elementary set theory. A compilation of definitions and theorems precedes each chapter --topics arranged as in Halmos' *Naive Set Theory*. Presumes a familiarity with algebraic structures --e.g., groups, rings, fields, vector spaces, etc. Reprint of the 1966 Van Nostrand edition. LCL

FOUNDATIONS, T*(17-18: 1, 2), L*. *Mathematical Logic*. J. Donald Monk. Grad. Texts in Math., V. 37. Springer-Verlag, 1976, x + 531 pp, $19.80. Clear, well-rounded presentation. Recursion theory (starting from Turing machines), logic, decidable and undecidable theories, model theory, and unusual logics. PJC

FOUNDATIONS, P. *Lecture Notes in Mathematics-521: Model Theoretic Algebra: Selected Topics.* Greg Cherlin. Springer-Verlag, 1976, iv + 234 pp, $9.50 (P). Lectures given at MIT and Heidelberg in 1974-5 with emphasis on transfer theorems and existentially complete structures in algebra. Necessary model theoretic background is sketched in Chapter 0. LCL

FOUNDATIONS. *The Structure and Dynamics of Theories.* Wolfgang Stegmüller. Springer-Verlag, 1976, xvii + 284 pp, $29.80. Highly sophisticated philosophy of science, with occasional resort to mathematical logic. Part I exposits J.D. Sneeds' *The Logical Structure of Mathematical Physics,* while Part II delves in depth into T.S. Kuhn's *The Structure of Scientific Revolutions* and the controversy it has generated. PJC

COMBINATORICS, P. *A Short Course on Error Correcting Codes.* N.J.A. Sloane. Springer-Verlag, 1975, 78 pp, $5.80 (P). A brief introduction to algebraic coding theory. Discusses linear codes, cyclic codes, Golay codes, the Macwilliams-Gleason theorems. Extensive bibliography; mentions some unsolved problems. SG

NUMBER THEORY, P. *Topics in Number Theory.* Ed: P. Turán. North-Holland, 1976, 456 pp, $47.95. A collection of some 35 papers on such topics as Diophantine approximations, L-series, multiplicative and additive functions, prime number problems, Gelfond-Baker methods, and quadratic forms. SG

NUMBER THEORY, T**(16-17: 1), S, P, L. *Modular Functions and Dirichlet Series in Number Theory.* Tom M. Apostol. Grad. Texts in Math., V. 41. Springer-Verlag, 1976, x + 198 pp, $14.80. This is the sequel to the author's *Introduction to Analytic Number Theory.* The emphasis is on elliptic and modular functions and their applications in number theory. A knowledge of basic complex variable theory is assumed. An excellent collection of problems is included in this well-written text. CEC

NUMBER THEORY, P. *Lecture Notes in Mathematics-536: Equations over Finite Fields, an Elementary Approach.* Wolfgang M. Schmidt. Springer-Verlag, 1976, ix + 267 pp, $10.30 (P). Stepanov's proof of Weil's theorem on curves over finite fields is given. This proof does not depend on algebraic geometry, but on methods from Diophantine equations. CEC

LINEAR ALGEBRA, T(13-14: 1). *A Geometric Introduction to Linear Algebra, Second Edition.* Daniel Pedoe. Chelsea, 1976, xi + 226 pp, $7.50. In essence a reprint (with minor changes) of the 1963 first edition. This text presents a relatively classical version of linear algebra (no pivoting, no eigenvalues, no linear programming) with a great deal of low dimensional Euclidean motivation. JAS

LINEAR ALGEBRA, S(14). *Matrices for Scientists and Engineers.* W.W. Bell. Van N-Rein, 1975, xii + 229 pp, $21.50. Treatment of matrices through numerical solutions of eigenvalues. Lengthy examples. Seems better suited for use as a reference text than as a classroom text. LLK

LINEAR ALGEBRA, T(13-18: 1), L. *Mathematical Tools for Applied Multivariate Analysis.* Paul E. Green, J. Douglas Carroll. Acad Pr, 1976, xiii + 376 pp, $24.50. Succinct treatment of the transformational geometry and matrix algebra relevant for the study of multiple regression, principal components analysis, and multivariate discriminant analysis. A pragmatic approach (no formal proofs) for students with no prior matrix algebra but some statistics. Appendices on differential calculus (including Lagrange multipliers) and generalized inverses. Designed for and well suited to students in the behavioral sciences. PJC

ALGEBRA, T*(16), L*. *Groups, Representations, and Characters.* Victor E. Hill. Hafner Pr, 1975, x + 181 pp, $12.95. A text on group representations for undergraduates, including students of chemistry and physics as well as mathematics. Begins with a review of group theory, then moves to representations (definitions, examples, regular and irreducible representations, representations of Abelian groups), characters (orthogonality relations, character tables, reducible characters), and conclude with special topics (real characters, induced representations and characters, space groups, linear groups that arise in physics). Deserves serious consideration as a text for a second course in algebra. SG

ALGEBRA, L. *Algebraic Extensions of Fields, Second Edition.* Paul J. McCarthy. Chelsea, 1976, ix + 166 pp, $8.50. A reprint of the first edition; see extended review, March 1968. SG

ALGEBRA, P. *Discrete Subgroups of Lie Groups and Applications to Moduli.* Walter L. Baily, Jr., *et al.* Oxford U Pr, 1975, 348 pp, $13.75 (P). Papers presented at the international colloquium in Bombay in January 1973. JAS

ALGEBRA, T(18: 2, 3), P, L*. *Algebraic Theories.* Ernest G. Manes. Grad. Texts in Math., V. 26. Springer-Verlag, 1976, 356 pp, $22.80. Exposition of Lawvere's notion of an algebraic theory as a nice category, along with many examples, exercises, and applications in topological dynamics and automata theory. Excellent book, good second course in algebra if first course was slightly categorical. PJM

ALGEBRA, P. *Lecture Notes in Mathematics-528: Ordinary and Modular Representations of Chevalley Groups.* James E. Humphreys. Springer-Verlag, 1976, 126 pp, $7.40 (P). Representations for subgroups of the general linear group over a finite field. PJM

ALGEBRA, T(16-17: 1), S, P. *An Introduction to Semigroup Theory.* J.M. Howie. Acad Pr, 1976, x + 272 pp, £9.80. Assumes little knowledge of abstract algebra but considerable mathematical maturity. Far less comprehensive than Clifford and Preston, but contains many more recent results, chiefly on regular semigroups. Problems, and an extensive bibliography. JD-B

ALGEBRA, P. *Dimension Theory for Nonsingular Injective Modules.* K.R. Goodearl, Ann K. Boyle. Memoirs No. 177. AMS, 1976, viii + 112 pp, $7.20 (P). Development of a dimension theory for nonsingular modules over regular, right self-injective rings. PJM

ALGEBRA, P. *Finite Free Resolutions.* D.G. Northcott. Tracts in Math., No. 71. Cambridge U Pr, 1976, xii + 271 pp, $29.50. A "self-contained" introduction to the study of modules possessing free resolutions of finite length. Includes basic material on matrices, localization, free modules as well as discussion of fitting to MacRae invariants, stability, latent nonzerodivisors, theory of grade, multiplicative structures. Includes some exercises with solutions. SG

CALCULUS, P. *Analysis.* F. Wille. Teubner, Stuttgart, 1976, 336 pp, DM 29 (P). One of a series of texts intended for prospective mathematics teachers in German academic high schools. Abstract and sophisticated, but well motivated. JD-B

CALCULUS, T(13: 1). *Elements of Calculus with Contemporary Applications.* Marcus M. McWaters, James H. Reed. HarBrace J, 1976, xi + 484 pp, $12.95. A text intended for a short course at an intuitive level. Covers algebra and geometry review, sequences, single and multivariable differential calculus for polynomials, logs and exponential functions, as well as integration of such functions. There is a concerted effort to keep the material from looking difficult: more artwork than most books as well as a large number of worked examples. TAV

REAL ANALYSIS, T?(16-17: 1), S, L. *A Brief Course in the Theory of Functions of a Real Variable (An Introduction to the Theory of the Integral).* B.Z. Vulikh. Trans: I.G. Volosova. MIR, 1976, 356 pp. A concise, easily understandable treatment of Lebesgue measure and integration, including discussions of L, L^p, and the Radon-Nikodym theorem. There are no exercises, making this a doubtful choice for a text. TAV

REAL ANALYSIS, P. *Lecture Notes in Mathematics-508: Regularly Varying Functions.* Eugene Seneta. Springer-Verlag, 1976, v + 112 pp, $7.40 (P). Study of functions satisfying $\lim f(\lambda x)/f(x) = \lambda^p$ as $x \to \infty$. Characterizations, representations. Contains proof of fact that measurable solutions of $f(x+y) = f(x) + f(y)$ are continuous. RBK

DIFFERENTIAL EQUATIONS. *Equations aux dérivées partielles.* Charles Blanc. Int. Ser. Num. Math., V. 34. Birkhauser, 1976, 136 pp, sFr. 28 (P). Based on a course for engineers. Considers equations arising from the study of evolution, elasticity, and vibrations among others. SG

DIFFERENTIAL EQUATIONS, P. *Inequalities in Mechanics and Physics.* G. Duvaut, J.L. Lions. Trans: C.W. John. Grund. math. Wissenschaften, B. 219. Springer-Verlag, 1976, xvi + 397 pp, $40.20. Partial differential inequalities arising in stationary problems and problems of evolution in classical physics and mechanics. Seven independent chapters. 200 references. Does not treat singular perturbations or numerical solutions of inequalities of evolution. DFA

DIFFERENTIAL EQUATIONS, T(17: 1), S, P. *Introduction to Partial Differential Equations.* Gerald B. Folland. Princeton U Pr, 1976, vi + 349 pp, $8.50 (P). An introduction to partial differential equations which assumes previous exposure to real and complex variable theory and the basic theorems of advanced calculus. Definitely a graduate level approach. There are not enough exercises. CEC

NUMERICAL ANALYSIS, S(16-17), P. *Vorlesungen über Numerische Mathematik.* Heinz Rutishauser. [illegible], B. 50 & 57. Birkhauser, 1976. *Band 1: Gleichungssysteme, Interpolation und Approximation,* 164 pp, sFr. 40; *Band 2: Differentialgleichungen und Eigenwertprobleme,* 228 pp, sFr. 48. Lectures given by the author at the Swiss Federal Institute of Technology, prepared for publication after his death by M. Gutknecht, with the help of P. Läuchli and H.-R. Schwarz. Vol. 1 deals with systems of equations, interpolation and approximation; Vol. 2 with differential equations and characteristic values. JD-B

NUMERICAL ANALYSIS, T(17), S, P. *A Survey of Numerical Methods for the Solution of Fredholm Integral Equations of the Second Kind.* Kendall E. Atkinson. SIAM, 1976, vii + 230 pp, $17.50 (P). Introduces foundations from functional analysis. Studies projection, degenerate kernel, collocation, Galenkin, Nyström methods and their variants. Includes error analysis, examples, and computer codes. RWN

NUMERICAL ANALYSIS, P. *Sparse Matrix Computations.* Ed: James R. Bunch, Donald J. Rose. Acad Pr, 1976, xi + 453 pp, $15. 26 papers by 36 authors, including the editors, presented at a symposium held at Argonne in September 1975. Categorization: sparse elimination, eigenvalue calculation, optimization, math software, pde's, and applications. RWN

NUMERICAL ANALYSIS, S(16-17), P. *Numerical Solution of Two Point Boundary Value Problems.* Herbert B. Keller. SIAM, 1976, viii + 61 pp, $4.80 (P). Based on ten lectures given at a regional conference held at Texas Tech in 1975. Shooting methods, finite difference methods, eigenvalue problems, singular problems. A very good, though brief, survey. References. RWN

NUMERICAL ANALYSIS, S(16-17), P. *Methods of Numerical Mathematics.* G.I. Marchuk. Trans: Jiri Ruzicka. Appl. of Math., No. 2. Springer-Verlag, 1975, xii + 316 pp, $29.80. Translated from the Russian. Based on the author's lecture notes. Considers broad classes of problems, e.g., stationary, non-stationary and inverse problems, especially those of interest in mathematical physics, e.g., solutions to pde's and transport problems. RWN

FUNCTIONAL ANALYSIS, T(18), P, L*. *Nonlinear Operators and Differential Equations in Banach Spaces.* Robert H. Martin, Jr. Wiley, 1976, xi + 440 pp, $27.50. Relates two areas of study: (1) spaces of nonlinear operators, elementary spectral theory, and fixed point theorems; (2) ordinary differential equations, linear semigroup theory, and semilinear differential equations. Provides preliminary material, chapters devoted exclusively to examples, extensive problem sections, references, notes and remarks. Attractive and carefully written. DFA

FUNCTIONAL ANALYSIS, P. *Lecture Notes in Mathematics-531: The Topology of Uniform Convergence on Order-Bounded Sets.* Yau-Chuen Wong. Springer-Verlag, 1976, vi + 163 pp, $7.40 (P). Monograph on one of the possible topologies of an ordered topological vector space. PJM

FUNCTIONAL ANALYSIS, S(16-17), P, L. *Theory of Distributions: The Sequential Approach.* Piotr Antosik, Jan Mikusiński, Roman Sikorski. PWN, 1973, xiv + 273 pp, $19.50. An important exposition of the theory of distributions. Distributions are defined as equivalence classes of "fundamental sequences" of functions. (A sequence is fundamental if for some k it is the sequence of kth derivatives of an almost uniformly convergent sequence of functions.) Some readers may find this approach more natural than that of Sobolev or Schwartz. The definition is in the spirit of Cantor's construction of the reals. J. Mikusiński's earlier operational calculus involves a construction similar to the construction of the rationals from the integers. RBK

FUNCTIONAL ANALYSIS, S(18), P. *Integral Equations--A Reference Text.* P.P. Zabreyko, *et al.* Noordhoff, 1975, xix + 443 pp, Dfl. 120. Presents main results on equations solvable in "closed form", classical Fredholm theory and its generalization to equations with compact operators, Hilbert-Schmidt theory, equations with nonnegative kernels, one- and multi-dimensional singular equations, equations of convolution type. Also examines applications to physics and some nonlinear equations. Cites more detailed treatments. DFA

FUNCTIONAL ANALYSIS, T(17-18: 1), S*, P. *Embeddings and Extensions in Analysis.* J.H. Wells, L.R. Williams. Ergebnisse der Math., B. 84. Springer-Verlag, 1975, vii + 107 pp, $14.40. Early in the book, the authors identify the dominant theme for each of the following two types of problems: (1) isometrically embed metric spaces in Hilbert spaces and (2) extend a given contraction map (defined on a subset) to the entire metric space. A unified treatment is then presented, focusing on a common class of quadratic forms. I-CH

FUNCTIONAL ANALYSIS, P. *Non-Commutative Spectral Theory for Affine Function Spaces on Convex Sets.* Erik M. Alfsen, Frederic W. Shultz. Memoirs No. 172. AMS, 1976, xii + 120 pp, $7.60 (P). Applies to spaces of self-adjoint elements of a von Neumann algebra, affine functions on a (Choquet) simplex, continuous affine functions on a rotund compact convex set such as unit ball of L^p, $1 < p < \infty$. Functional calculus developed. New proofs of old results are more geometrical. RBK

OPTIMIZATION, P. *Lecture Notes in Economics and Mathematical Systems-129: Komplementaritäts-und Fixpunktalgorithmen in der mathematischen Programmierung, Spieltheorie und Ökonomie.* Hans-Jakob Lüthi. Springer-Verlag, 1976, 145 pp, $7.40 (P). An account, for specialists, of recent work on complementarity and fixed-point algorithms. JD-B

OPTIMIZATION, T(18), P. *Applied Functional Analysis.* A.V. Balakrishnan. Appl. of Math., No. 3. Springer-Verlag, 1976, x + 309 pp, $19.80. Revised and enlarged version of the author's *Introduction to Optimization Theory in a Hilbert Space.* Basic properties of Hilbert spaces, convex sets and convex programming, functions, transformations and operators, semigroups of linear operators, optimal control theory, probability measures on a Hilbert space. Problems. Assumes good background in real and complex analysis. DFA

OPTIMIZATION, S*(17-18), P. *A Contribution to Theory and Practice of Nonlinear Parameter Optimization.* Ph. Th. Stol. Pudoc (US Distr: ISBS), 1975, 197 pp, $19 (P). Essentially a reproduction of the author's 1975 doctoral thesis. Studies the least squares nonlinear parameter optimization using differential geometry techniques. Lays out the technical procedure: project the search-path (on the fitting surface) onto the tangent plane; correct the direction-deviation by optimizing angles and associate the convergence rate with the search path curvature. Resourceful, well organized and implemented (by FORTRAN programs). Definitely needs an index. I-CH

OPTIMIZATION, T(17-18: 1), P. *Optimization Methods.* Henning Tolle. Springer-Verlag, 1975, xiv + 226 pp, $19.80 (P). Based on Caratheodory's approach to the calculus of variations, this book presents various methods (e.g., Pontryagin's maximum principle, the gradient method and Bellman's dynamic programming method, among others), available only since 1950. Draws examples almost exclusively from space flight technology, due to author's earlier work in space industry. I-CH

OPTIMIZATION, P. *Optimization Over Leontief Substitution Systems.* Ed: Gary J. Koehler, Andrew B. Whinston, Gordon P. Wright. North-Holland, 1975, x + 221 pp, $18.75. The development of "efficient methods for solving large scale linear programs" based on matrix iterative methods. Includes extensions, applications and examples as well as proofs. Bibliography. RWN

ANALYSIS, P. *Lecture Notes in Mathematics-518: Séminaire de Théorie du Potentiel Paris 1972-1974.* F. Hirsch, G. Mokobodzki. Springer-Verlag, 1976, vi + 275 pp, $10.30 (P).

ANALYSIS, P. *Ergodic Theory and Topological Dynamics.* James R. Brown. Pure and Appl. Math., V. 70. Acad Pr, 1976, x + 190 pp, $19.50. A survey of ergodic theory and dynamical systems. Begins with a discussion of the various ergodic theorems, topological dynamics, affine transformations of compact abelian groups, entropy, and ends with an exposition of Ornstein's isomorphism theorem. Many interesting exercises are included. SG

ANALYSIS, T*(15: 1, 2). *Advanced Calculus.* Kenneth Rogers. Merrill, 1976, ix + 365 pp, $16.95. Assumes elementary linear algebra. Topology, continuity, differentiability, (Riemann) integration on the line and, separately, in n-space. Infinite series of constants and of functions. Curves and surfaces. Problem set with many challenging entries follows each section. Also suitable for a one-semester, one-variable course. Attractively illustrated. DFA

ANALYSIS, P. *Lecture Notes in Mathematics-494: Order and Potential Resolvent Families of Kernels.* Aurel Cornea, Gabriela Licea. Springer-Verlag, 1975, 154 pp, $7.40 (P). Beginning with the idea of a σ-lattice cone, the authors develop a unified theory of excessive functions and excessive measures. The key results are (1) Hunt's theorem for the existence of a resolvent and (2) all right continuous supermartingales may be regarded as excessive elements with respect to a convenient resolvent family of kernels. TAV

ANALYSIS, P, *Géométrie Symplectique et Physique Mathématique*. Ed: Jean-Marie Souriau. Cent Nat Rech Scient, 1975, 425 pp, 150,00 F. 26 papers on many aspects of symplectic geometry and its applications in mathematical physics, from a conference at Aix-en-Provence in June 1974. 21 are in English, 4 in French, 1 (by Klingenberg) in German. DFA

ANALYSIS, P, *Lecture Notes in Mathematics-532: Théorie Ergodique*. Ed: J.-P. Conze, M.S. Keane. Springer-Verlag, 1976, viii + 227 pp, $9.50 (P). A wide ranging set of papers presented at the "Journées Ergodiques" in Rennes during the years 1973 and 1974. JAS

ANALYSIS, P, *Lecture Notes in Mathematics-527: Ergodic Theory on Compact Spaces*. Manfred Denker, Christian Grillenberger, Karl Sigmund. Springer-Verlag, 1976, iv + 360 pp, $13.20 (P). A broad introduction to topological ergodic theory on compact spaces. Among the topics: invariant measures, shifts, measure theoretic and topological entropy, decompositions, generators, stability, aperiodic transformations. SG

GEOMETRY, L*, *Fantasy & Symmetry, The Periodic Drawings of M.C. Escher*. Caroline H. MacGillavry. Abrams, 1976, xi + 84 pp, $15. Reprint of a 1965 volume published for the Inter. Union of Crystollography. Each of the 41 plates is accompanied by a discussion by the crystollographer-editor of the symmetry patterns illustrated in the plates. LAS

GEOMETRY, P, *Der barycentrisch Calcul*. August Ferdinand Möbius. Georg Olms, 1976, 454 pp, DM 34. Almost a facsimile of the original Leipzig edition of 1827. JD-B

GEOMETRY, T*(16-17: 1, 2), S, P, L, *Introduction to Finite Geometries*. F. Kárteszi. North-Holland, 1976, xiii + 266 pp, $39.75. A general introduction to the basic concepts of finite geometry along with detailed discussions of Galois geometries. Group theoretical and combinatorial methods, as well as methods using the coordinate system are discussed. A good collection of problems is included. This would make an excellent textbook for wealthy students. CEC

GEOMETRY, T(13-14: 1), S, *Grundlagen der Geometrie*. Rolf Nevanlinna, Paul Edwin Kustaanheimo. Math. Reihe, B. 43. Birkhauser, 1976, 135 pp, sFr. 34. Text is divided into two parts: affine geometry of the plane and finite geometry. First part starts with an axiomatic approach using the point and line as basic elements. Later contains material on vector algebra, affine coordinate systems and transformations, and different types of geometries. The second part discusses incidence, order, and congruence in geometries with only a finite number of points and lines. Bibliography. Index. RJA

TOPOLOGY, P, *Algebra, Topology, and Category Theory: A Collection of Papers in Honor of Samuel Eilenberg*. Ed: Alex Heller, Myles Tierney. Acad Pr, 1976, xi + 225 pp, $26.50. Papers representative of the wide range of areas in which S. Eilenberg has made contributions. JAS

TOPOLOGY, S(16-17), L, *Introduction a la Théorie des Surfaces de Riemann*. J. Guenot, R. Narasimhan. L'Enseignement Math, 1976, 214 pp, frs. 35 (P). Course in Riemann surfaces from definition of manifold through Riemann-Roch theorem. No exercises. Presents results with as little algebraic topology as possible. PJM

TOPOLOGY, P, *Anti-Invariant Submanifolds*. Kentaro Yano, Masahiro Kon. Pure and Appl. Math., V. 21. Dekker, 1976, vii + 183 pp, $19.75 (P). A sub-manifold of a (complex) manifold is anti-invariant if the image of the tangent space under the complex structure is contained in the normal space. Monograph covers recent work on anti-invariant submanifolds of Kahlerian and Sasakian manifolds. PJM

TOPOLOGY, T(16-17), L*, *Knots and Links*. Dale Rolfsen. Publish or Perish, 1976, 439 pp, $15 (P). A nice introduction to modern knot theory. Lots of good pictures including an appendix of all possible knots and links of up to ten crossings (nine for links). Exercises, good bibliography. Prerequisites: point-set and algebraic topology up to fundamental group. Second half of book needs some PL-topology. PJM

TOPOLOGY, P, *Lecture Notes in Mathematics-540: Categorical Topology*. Ed: E. Binz, H. Herrlich. Springer-Verlag, 1976, xv + 719 pp, $20.10 (P). Proceedings of the conference held at Mannheim in July 1975. JAS

TOPOLOGY, P, *Continuous Cohomology of Spaces with Two Topologies*. Mark Alan Mostow. Memoirs No. 175. AMS, 1976, x + 142 pp, $8 (P). If a space has two topologies, X and X' with X' weaker, one gets a cohomology by looking at real cochains of X' continuous in X. The author looks at axioms for this sort of cohomology, examples, and applications. PJM

TOPOLOGY, T(17-18: 1), S, P, *Differential Topology*. Morris W. Hirsch. Grad. Texts in Math., V. 33. Springer-Verlag, 1976, x + 221 pp, $14.80. Prerequisites are some introductory analysis and general topology (contained in an appendix). Author believes that homology and homotopy are better appreciated after a study of the present material, whose exposition avoids algebraic methods. Chapters on transversality, tubular neighborhoods, degrees of maps, the Euler characteristic, Morse theory, cobordism, etc. Exercises. Bibliography. Index. RJA

PROBABILITY, P, *Lecture Notes in Mathematics-526: Probability in Banach Spaces*. Ed: A. Beck. Springer-Verlag, 1976, vi + 290 pp, $11.50 (P). Proceedings of an international conference on probability in Banach spaces held at Oberwolfach in July 1975. JAS

PROBABILITY, T(14-15: 1), L, *Basic Probability Theory and Applications*. Ramakant Khazanie. Goodyear, 1976, xi + 516 pp, $15.95. A one semester course in probability, this text provides a natural lead-in to a course in either mathematical statistics or stochastic processes. The treatment is honest, e.g., the author defines random variables in terms of Borel sets, and very intuitive as well. Covers multivariate random variables, generating functions and standard limit theorems. A strong point: numerous, well chosen realistic examples and exercises. TAV

PROBABILITY, T(14-16: 1, 2). *Probability.* Peter Whittle. Wiley, 1976, 239 pp, $7.50 (P). A reprinting of a 1970 Penguin original (TR, November 1970; ER, October 1971). LAS

PROBABILITY, T(18: 1), P. *Lecture Notes in Mathematics-511: Séminaire de Probabilités X Université de Strasbourg.* Ed: P.A. Meyer. Springer-Verlag, 1976, vi + 593 pp, $18.10 (P). The lectures from the probability theory seminar at the University of Strasbourg in 1974-1975 together with a systematic presentation of the theory of stochastic integrals. JAS

PROBABILITY, T*(18), P*, L*. *Denumerable Markov Chains, Second Edition.* John G. Kemeny, J. Laurie Snell, Anthony W. Knapp. Grad. Texts in Math., V. 40. Springer-Verlag, 1976, xii + 484 pp, $16.80. Differs from the original (1966) edition by the addition of a chapter on Markov random fields, as well as a section of additional notes indicating the developments in the area over the last decade with appropriate bibliographic references. TAV

PROBABILITY, S(16-17), P. *Wahrscheinlichkeits-Theorie, Eine Einführung.* A.A. Borowkow. Birkhauser, 1976, xi + 264 pp, sFr. 38. A translation, with minor additions and corrections, of a Russian introduction to probability theory. The necessary measure and integration theory are treated in appendices. No problems or bibliography. JD-B

PROBABILITY, P. *Lecture Notes in Mathematics-516: Boundary Theory for Symmetric Markov Processes.* Martin L. Silverstein. Springer-Verlag, 1976, xvi + 313 pp, $12.30 (P). An extension of the author's previous monograph (Lecture Notes #426) on symmetric Markov processes. In this work attention is paid to the construction of symmetric processes on the boundary of the original process, and the nature of the associated Dirichlet space. TAV

PROBABILITY, P. *Lecture Notes in Mathematics-539: Ecole d'Eté de Probabilités de Saint-Flour V-1975.* A. Badrikian, J.F.C. Kingman, J. Kuelbs. Springer-Verlag, 1976, ix + 314 pp, $12.30 (P). Three major papers: Prolegomenes au calcul des probabilities dans les Banach by A. Badrikian; Subadditive processes by J.F.C. Kingman; The law of the iterated logarithm and related strong convergence theorems for Banach space valued random variables by J. Kuelbs. JAS

PROBABILITY, T(18: 1), S, P. *Lecture Notes in Mathematics-534: Random Fields.* Chris Preston. Springer-Verlag, 1976, 200 pp, $9.50 (P). These notes are concerned with the properties of equilibrium states defined in terms of conditional probability. The approach is fairly abstract and uses the language and basic techniques of probability theory. CEC

STATISTICS, T(15-16: 1). *Grundinhalte der Statistik.* Foster Lloyd Brown, Jimmy Y. Amos, Oscar G. Mink. Hermann Schroedel, 1976, 160 pp, DM 24 (P). A translation of the second American edition, which was called *Statistical Concepts: A Basic Program,* published by Harper and Row. JD-B

STATISTICS, T(16-18: 1, 2), S, P, L. *Theory and Application of the Linear Model.* Franklin A. Graybill. Duxbury Pr, 1976, xiv + 704 pp, $22.50. Analysis of variance, correlation and regression, and design of experiments. Largely "normal theory." Includes review of matrix algebra, basic statistics and normal distribution theory. FLW

STATISTICS, S(15-17), P, L. *Applications of Statistics to Industrial Experimentation.* Cuthbert Daniel. Wiley, 1976, xvi + 294 pp, $17.50. "Statistical aids available for the planning of industrial experiments" and the "analysis and interpretation of the data collected from such experiments." "Cookbook or research monograph?" "A bit of each." FLW

STATISTICS, T(16-18: 1, 2), S, P, L. *Dynamic Stochastic Models from Empirical Data.* R.L. Kashyap, A. Ramachandra Rao. Math. in Sci. and Eng., V. 122. Acad Pr, 1976, xvi + 334 pp, $34.50. The construction of models for empirical time series, the "development of various approaches to comparison of different classes and subsequent validation of the final models", and "relatively standard topics such as parameter estimation methods and estimability." FLW

STATISTICS, P. *Statistische Tafeln zur multivariaten Analysis.* Heniz Kres. Springer-Verlag, 1975, xviii + 431 pp, $19.70 (P). Twenty-six tables, many of them reprinted from journals, with suggestions for their use and references to the literature. JD-B

STATISTICS, T(15-17: 1-3), S, P, L. *Time Series Analysis, Forecasting and Control, Revised Edition.* George E.P. Box, Gwilym M. Jenkins. Holden-Day, 1976, xxi + 575 pp, $32.95. Stochastic model building, identification, estimation, and checking. Transfer function models. An outline of computer programs. This edition (First edition, TR February 1971) also has a new section containing exercises--many of them based on real data. FLW

COMPUTER SCIENCE, S, P, L. *High-Level Language Computer Architecture.* Ed: Yaohan Chu. Acad Pr, 1975, xii + 273 pp, $29.50. Seven tutorial articles and case studies on designing computer architecture to facilitate the writing of software systems and compilers in high-level languages. Authors are Carlson, Doran, Laliotis, Bloom, Robinet and the editor. RWN

APPLICATIONS (PHYSICS), P. *Studies in Mathematical Physics: Essays in Honor of Valentine Bargmann.* Ed: E.H. Lieb, B. Simon, A.S. Wightman. Princeton U Pr, 1976, xii + 460 pp, $26.50; $10 (P).

Reviewers Whose Initials Appear Above

Richard J. Allen, St. Olaf; David F. Appleyard, Carleton; Paul J. Campbell, St. Olaf; Clifton E. Corzatt, St. Olaf; John Dyer-Bennet, Carleton; Steven Galovich, Carleton; Ih-Ching Hsu, St. Olaf; Lorraine L. Keller, St. Olaf; Roger B. Kirchner, Carleton; Loren C. Larson, St. Olaf; Pierre J. Malraison, Carleton; R.W. Nau, Carleton; J. Arthur Seebach, Jr., St. Olaf; Lynn A. Steen, St. Olaf; T.A. Vessey, St. Olaf; Frank L. Wolf, Carleton.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, NW, Washington, D. C. 20036. Items must be submitted at least five months before publication can take place.

PERSONAL ITEMS

Marshall University: Assistant Professor Elizabeth Czompo and Associate Professor Layton Thompson retired in May 1976.

Rensselaer Polytechnic Institute: Professor Dis Maly has retired with the title of Professor Emeritus; Dr. B. F. Caviness, Illinois Institute of Technology, has been appointed Associate Professor.

Professor Yen-Mow Lynn, University of Maryland, Baltimore County, has been appointed Chairman of the Department of Mathematics.

Dr. Albert H. Lightstone, Queen's University, Ontario, Canada, died on March 29, 1976, at the age of 49. He was a member of the Association for fourteen years.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

MAY MEETING OF THE OHIO SECTION

The Ohio Section of MAA held its annual Spring meeting at Youngstown State University, Youngstown, May 7 and 8, 1976. Approximately one hundred and twenty people were in attendance. Section Chairman R. G. Laatsch presided; R. S. Varga was the Program Chairman. The theme for the meeting centered upon "Complex Function Theory and Applications."

Invited addresses included: "Geometric Convergence of Rational Functions to Analytic Functions in Unbounded Domains", by E. B. Saff, University of South Florida, Tampa; and "A Tribute to Arnold E. Ross on His Retirement", by H. J. Zassenhaus, Ohio State University. R. G. Laatsch (Miami University) presented the retiring Chairman's address, "Scrambled Infinite Dimensional Convex Eggs."

The following contributed papers were presented:

A Factorization for metacyclic groups, by D. C. Buchthal, University of Akron.

On the automorphism group of a strongly connected automata, by J. J. Buoni, Youngstown State University.

The minimum modulus of polynomials with ± 1 *as coefficients*, by D. Eustice, Ohio State University.

Extreme points in certain linear spaces, by E. P. Merkes, University of Cincinnati.

A brief survey of quasi-conformal mappings, by M. B. Osgood, student, Carnegie-Mellon University.

Musing about means, by L. D. Rodabaugh, University of Akron.

Regular expressions over a semiring, by E. S. Santos, Youngstown State University.

Lebesgue constants of (f, d_n)*-summability*, by R. A. Shoop, Kent State University.

Reciprocal sets of vectors in n-space, by G. L. Szoke, University of Akron.

Rational approximations of e^{-z} *and applications to the numerical solution of heat-conduction problems*, by R. S. Varga, Kent State University.

Swap sessions included: "Metrication — Some Implications and Applications", led by L. Ross, University of Akron; and "Meeting of Chairmen of State-Supported Schools," led by J. Landin, Ohio State University. Also, the agenda included the annual Business Meeting of the Section, the meeting of the Executive Committee, and the meeting of ad hoc committees: Committee on Cooperation among Colleges and Universities, Committee on Curriculum, and Committee on Teacher Training and Certification.

Section officers for academic year 1976–77 are: J. A. Murtha, Marietta College, Chairman; W. H. Beyer, University of Akron, Chairman-Elect; G. Mavrigian, Youngstown State University, Secretary-Treasurer; M. Wetzel, Denison University, Program Chairman; J. H. Carney, Lorain County Community College, and C. A. Long, Bowling Green State University, Program Committee. R. L. Wilson, Ohio Wesleyan University, serves as Sectional Governor.

GUS MAVRIGIAN, *Secretary-Treasurer*

REPORT OF THE TREASURER FOR THE YEAR 1975

Herewith is a summary of the report of the Treasurer of the Association for the year 1975. In this summary, all entries have been rounded to the nearest dollar; therefore sums of entries may differ from the entered total. The full report has been approved by the Finance Committee and accepted by a vote of the Board of Governors. Any member of the Association who wishes to have a copy of the full report may obtain one by writing to the Washington Office of the Association.

ASSETS	Dec. 31, 1975
Cash	$ 87,070
Securities (at cost), unrestricted	244,462
Securities (at cost), restricted	151,010
Accounts Receivable	49,293
Furniture and Equipment	16,426
Prepaid Expenses	29,257
Two-Year College Mathematics Journal	20,800
Total Assets	$ 598,318
LIABILITIES	
Accounts Payable	$ 21,718
Unearned Income	
Dues and Subscriptions	324,643
Other	17,728
Prindle–Weber–Schmidt Advertising Account	23,150
NSF Fund	15,800
High School Contest Fund	24,654
Total Liabilities	$ 427,693
Assets minus Liabilities (Net Worth, including restricted funds).	$ 170,625

OPERATING INCOME	
Dues	$ 297,605
Publications	295,610
Dividends and Interest	22,322
Contributions	13,966
Registration Fees	10,090
Indirect Costs — Outside Agencies	13,034
Indirect Costs — High School Contest	6,000
Miscellaneous	4,357
Total Operating Income	$ 662,982

OPERATING EXPENDITURES	
Salaries	$ 238,440
Office Expenses	110,840
Publications	234,351
Travel and Meeting Expenses	42,133
Taxes and Fees	8,169
Dues and Contributions	13,700
Awards and Grants	12,667
Miscellaneous	5,528
Total Operating Expenditures	$ 665,827
Operating Income over (under) Operating Expenditures	$ (2,845)

LEONARD GILLMAN, *Treasurer*

CALENDAR OF FUTURE MEETINGS

Sixtieth Annual Meeting, St. Louis, Missouri, January 29-31, 1977.

Fifty-seventh Summer Meeting, University of Washington, August 14-16, 1977.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN, St. Francis College, Loretto, Pennsylvania, April 22-23, 1977.

FLORIDA, University of South Florida, Tampa, March 4-5, 1977.

ILLINOIS, Chicago Loop College, Chicago, May 6-7, 1977.

INDIANA

INTERMOUNTAIN

IOWA, Drake University, Des Moines, April 22-23, 1977.

KANSAS, Tabor College, Hillsboro, April 2, 1977.

KENTUCKY, early April. Deadline for papers 6 wks. bef. mtg.

LOUISIANA-MISSISSIPPI, University of New Orleans, Louisiana, February 25-26, 1977.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA, Saturday before Thanksgiving and last Saturday in April.

METROPOLITAN NEW YORK, Spring. Deadline for papers 2 wks. bef. mtg.

MICHIGAN, Eastern Michigan University, Ypsilanti, May 6-7, 1977.

MISSOURI, University of Missouri, St. Louis, April 29-30, 1977.

NEBRASKA, Nebraska Wesleyan University, Lincoln, April 15-16, 1977.

NEW JERSEY, early November and early May.

NORTH CENTRAL, North Hennepin Community College, Minneapolis, Minnesota, April 29-30, 1977.

NORTHEASTERN, Saturday after Thanksgiving, and third week in June in odd-numbered years.

NORTHERN CALIFORNIA, San Francisco State University, February 26, 1977.

OHIO

OKLAHOMA-ARKANSAS, Oral Roberts University, Tulsa, Oklahoma, April 1-2, 1977.

PACIFIC NORTHWEST, second Saturday in June. Deadline for papers 6 wks. bef. mtg.

PHILADELPHIA, Saturday before Thanksgiving.

ROCKY MOUNTAIN, Metropolitan State College, Denver, Colorado, April 29-30, 1977.

SEAWAY, first Saturday in November and Saturday in late April. Deadline for papers 6 wks. bef. mtg.

SOUTHEASTERN, University of Alabama, Huntsville, April 1-2, 1977.

SOUTHERN CALIFORNIA, Loyola Marymount University, Los Angeles, March 12, 1977.

SOUTHWESTERN, Phoenix College, Phoenix, Arizona, April 22-23, 1977.

TEXAS, Baylor University, Waco, April 1-2, 1977.

WISCONSIN, University of Wisconsin, Oshkosh, Spring 1977.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Denver, February 20-26, 1977.

AMERICAN MATHEMATICAL ASSOCIATION OF TWO YEAR COLLEGES

AMERICAN MATHEMATICAL SOCIETY, St. Louis, Missouri, January 27-30, 1977.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION, University of North Dakota, Grand Forks, June 13-16, 1977.

ASSOCIATION FOR COMPUTING MACHINERY, Olympic Hotel, Seattle, Washington, October 17-19, 1977.

ASSOCIATION FOR SYMBOLIC LOGIC, Chase Park Plaza Hotel, St. Louis, Missouri, January 27-28, 1977.

ASSOCIATION FOR WOMEN IN MATHEMATICS, St. Louis, Missouri, January 29-30, 1977.

CANADIAN SOCIETY FOR HISTORY AND PHILOSOPHY OF MATHEMATICS/SOCIÉTÉ CANADIENNE D'HISTOIRE ET DE PHILOSOPHIE DES MATHÉMATIQUES, Hamilton, Ontario, June 2, 1977

FIBONACCI ASSOCIATION

INSTITUTE OF MATHEMATICAL STATISTICS, Seattle, Washington, August 14-18, 1977.

MU ALPHA THETA

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Cincinnati, Ohio, April 20-23, 1977.

OPERATIONS RESEARCH SOCIETY OF AMERICA, San Francisco Hilton, May 9-11, 1977.

PI MU EPSILON

SCHOOL SCIENCE AND MATHEMATICS ASSOCIATION

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS

ESSENTIALS OF ALGEBRA FOR COLLEGE STUDENTS

Raymond A. Barnett, Merritt College

An intermediate algebra text with the unique feature of substantial supportive material from elementary algebra. The pedagogic standards established and so widely praised in all of Ray Barnett's other texts are here—clear exposition, numerous worked examples with matched problems and answers, over 3200 carefully selected and graded exercises, interesting real world applications, and a Keller-type supplement in the Instructor's Manual for individualized, self-paced instruction. An effective means for overcoming algebraic deficiencies and preparing your students for subsequent courses. 384 pages, $11.95; Instructor's Manual available.

McGraw-Hill texts
Prime factors in any course

ELEMENTARY FUNCTIONS

Carl B. Allendoerfer, Cletus O. Oakley, Haverford College, and **Donald R. Kerr, Jr.,** Indiana University

Providing a brief and uncomplicated approach to elementary functions, this text emphasizes those topics from algebra, trigonometry, and analytic geometry needed to begin the study of calculus. Special attention is paid to graphing. More than 250 completely worked out examples and 1000 graded problems are included. The informal language of the text and its open, inviting format make this an easy text for both you and your students to use. An added benefit for you—the accompanying manual discusses each chapter in enough detail to be of value to everyone from the beginning instructor to the seasoned teacher. 352 pages, $12.50; Instructor's Manual available.

COLLEGE ALGEBRA, Seventh Edition

Paul K. Rees, Louisiana State University, **Fred W. Sparks,** Texas Tech University, and **Charles Sparks Rees,** University of New Orleans

The level and scope of this popular text have been maintained in the new edition. Greater emphasis is on the function concept and development of a more intuitive approach. The 4000+ problems and exercises—twice the number offered in other texts—have been thoroughly revised. Attention has been paid to providing excellent coverage of review topics—the four fundamental operations, special products, and fractions—to insure students' understanding of basic concepts. New material on systems of linear inequalities and linear programming is also included. 512 pages, $13.50; Instructor's Manual available.

ON US

"*I really found them [the excellent examples and exercises] to be much better than many other books. . . . They illustrate the material beautifully. Further, they do reinforce the ideas presented.*" — Ralph D'Agostino, Boston University. The 320-page Study Guide provides additional exercises in a programmed format and complete solutions to problems answered in the text.

Patterns and Systems of Elementary Mathematics

Jonathan Knaupp, Lehi Smith, Paul Shoecraft, all of Arizona State University, and Gary Warkentin, Pacific College

About 400 pages / paper / Instructor's Manual / Just published

"*PSEM is the one text for teachers that I have seen that attempts to teach them as it is hoped they will teach.*" — Instructor, a California State University.

Patterns and Systems of Elementary Mathematics uses discovery activities and manipulative aids (included with the text) to help students learn and appreciate mathematics. Introductions and summaries to each chapter add to students' perspective.

The Instructor's Manual is an especially valuable adjunct to this program, containing detailed recommendations for teaching each topic, solutions, sample quizzes and tests for each chapter. Years of class-testing at the Arizona State University have proven this text highly effective.

Introductory Calculus with Applications, Second Edition

J. S. Ratti and M. N. Manougian, both of the University of South Florida

About 450 pages / Instructor's Manual
Just published

Ratti and Manougian add to their proven record of effective writing with the Second Edition of *Introductory Calculus with Applications.* They provide for a wider range of students with a more comprehensive review of algebra, 78% more exercises, more intuitive presentations of limits and the integral, and even more examples and applications, particularly to business, economics, and the calculator. The Second Edition can be used more flexibly for one or two semester courses because of an optional chapter on differential equations and sections on improper and multiple integrals.

To request examination copies for adoption consideration, contact your regional Houghton Mifflin sales office

Houghton Mifflin

Atlanta, GA 30324 • Dallas, TX 75235 • Geneva, IL 60134
Hopewell, NJ 08525 • Palo Alto, CA 94304 • Boston, MA 02107

THE MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.
Washington, DC 20036

THE AMERICAN
MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

AN OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 84 NUMBER 2

CODEN: AMMYAE

CONTENTS

(*Continued on inside cover*)

FEBRUARY 1977

NOTICE TO AUTHORS

Research papers per se are unsuitable; see statement of policy (Vol. 81, p. 1).

Please follow the format in current issues of the Monthly. Manuscripts must be legibly typewritten or reproduced from typewritten copy, double spaced with wide margins. Three copies should be submitted to the appropriate editor and one kept by the author as protection against loss. The author's full address *must* appear at the end of the manuscript.

Backlog: Main Articles 18 months, Math. Notes 17 months, Research Problems 21 months, Classroom Notes 16 months, Math. Education 17 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to R. P. Boas, Department of Mathematics, Northwestern University, Evanston, IL 60201; NOTES, etc.: to the corresponding Associate Editor; REPRINT PERMISSION: to Leonard Gillman, Mathematical Association of America, University of Texas, Austin, Texas 78712 (see also the copyright notice below); ADVERTISING CORRESPONDENCE: to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo, Buffalo, N.Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. Willcox, Mathematical Association of America, 1225 Connecticut Ave., N. W., Washington, D.C. 20036; BACK ISSUES: Contact P. and H. Bliss Co., Middletown, CT 06457.

Annual dues for members of the Association, including a subscription to the American Mathematical Monthly, are $18.00 for each of the first two years of membership and $22.00 thereafter. Student Membership is available with annual dues of $12.00. For nonmembers the subscription price is $25.00.

Published by the Association at Washington, D.C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D.C., and additional mailing offices.

PRINTED IN THE UNITED STATES OF AMERICA

AWARD FOR DISTINGUISHED SERVICE TO PROFESSOR VICTOR KLEE

Since the early 1960's the Mathematical Association of America has conferred an annual Award for Distinguished Service to Mathematics, an award that is made for "outstanding service to mathematics of such a character as to influence significantly the field of mathematics or mathematics education on a national scale." It would be difficult, I think, to find a member of the mathematical community who more abundantly qualifies than the recipient of this year's award, Vic Klee. In fact, when one looks over the record of his activities in mathematics and mathematical education, it is hard to see how one person could manage his time in such a way as to accomplish so much. I got some insight into this in Vic's case years ago when he used to say that he didn't mind doing chores like drying dishes, since that was not an intellectually demanding activity and he could at the same time be thinking of how to solve research problems in mathematics. I never found out whether any Klee theorems were proved at the expense of shattered crockery — you'd have to ask his wife, Bitsy, about that.

Victor Klee was born in 1925 in San Francisco and obtained his Bachelor of Arts degree at Pomona College in 1945, with high honors in mathematics and with a double major, in mathematics and chemistry. The first of his published papers, which now number well over 150, resulted from work done as an undergraduate on the Euler ϕ function. He then went for graduate work to the University of Virginia, where he held successively a DuPont Fellowship, an instructorship and an Atomic Energy Commission Predoctoral Fellowship. His doctorate at Virginia was obtained in 1949 under E. J. McShane. I had just joined the faculty there the previous fall, so I got to see Vic's thesis, which was on Convex Sets in Linear Spaces. I'll always remember the slightly unusual acknowledgment, a sentence of which read something like this: "The author expresses his special appreciation to Professor E. J. McShane, whose suggestion of a dissertation subject in the calculus of variations led obliquely to the present research on convex sets."

Vic's research work and qualities as a teacher were so impressive that he was invited to join the faculty of the University of Virginia where he remained for four years, except for one year of leave at the Institute for Advanced Study as a National Research Council Fellow. In 1952 work growing out of his dissertation, won for him the President and Visitors Research Prize in a competition open to faculty members in all fields at the University. It was rare enough for this prize to be won for work in mathematics and even rarer for it to be won by a person so young. In 1954 Vic joined the faculty of the University of Washington in Seattle, where by 1957 he had advanced to the rank of full professor and where he has remained since, except for leaves of absence. Within the period 1956–1960 he held both a Sloan Foundation Research Fellowship and a National Science Foundation Senior Postdoctoral Fellowship, spending the last two years of that period at the University of Copenhagen. He spent the major part of 1972 as a consultant at the Thomas J. Watson Research Center of the International Business Machines Corporation and the academic year 1975–76 at the Center for Advanced Study in the Behavioral Sciences. During his years at the University of Washington he has had around twenty PhD students, who now occupy faculty posts in universities and colleges scattered throughout the United States and Canada. Through these, through his own researches, and through his numerous invited lectures at regional, national and international conferences, Klee has certainly had a significant and widespread influence on mathematical research and education.

What is perhaps equally important for the present Award, Klee has a truly impressive record of service to the mathematical professional community and to governmental and industrial organizations in this country. We should mention first his presidency of this Association in 1971–72, but he has held many other important posts as well. He has served as Associate Secretary and as a member of the Council and of the Executive Committee of the American Mathematical Society. He has been a national Sigma Xi lecturer. He has been a member of the Council of the Society for Industrial and

VICTOR KLEE

Applied Mathematics and of both the Council and the Board of Trustees of the Conference Board of the Mathematical Sciences. He has served on the Advisory Panel for Mathematical Sciences of the National Science Foundation and has organized a symposium on applied combinatorics for the Office of Naval Research. He has won the L. R. Ford award for an expository paper on convex sets and was featured in an expository film on unsolved problems of geometry that won honorable mention at the National Educational Film Festival. In addition to his work for IBM already mentioned, Klee has been a consultant to the National Security Agency, the Boeing Scientific Research Laboratories, the RAND Corporation, and the DuPont Company. Most recently, this past spring, he was a member of the U.S. delegation of nine mathematicians that visited the Peoples Republic of China.

It would not do to close without touching on the personal qualities of the man we are honoring today. Those of us who know Vic Klee well especially value his integrity, unpretentiousness and quiet common sense; his friendliness, warmth and good humor; and his delight in a good joke and even a bad pun. I'll give you just one example, in which I was personally involved. One day at the University of Virginia, Vic came in briskly and out of the blue said to me, "Tru, do you have lots of hair on your chest? Puzzled but unwary, I replied, "Only the usual amount, I guess. Why do you ask?" Vic said, "Oh, I was just wondering whether it would be appropriate to call you Hairy Truman!" Now that the memory of that occasion has mellowed, it is a signal honor and pleasure for me to welcome Vic Klee, on behalf of the Mathematical Association of America, to the ranks of those who have won and richly deserved its Award for Distinguished Service to Mathematics.

TRUMAN BOTTS

ERROR-CORRECTING CODES AND INVARIANT THEORY: NEW APPLICATIONS OF A NINETEENTH-CENTURY TECHNIQUE

N. J. A. SLOANE

Abstract. An unfashionable nineteenth century technique, invariant theory, has recently been used to study error-correcting codes. This technique is potentially of much wider application, is very powerful, often produces startling results, and (not least) is fun to use.

I. INTRODUCTION

It will be best to begin with an example, showing how invariant theory is used to solve a typical problem. Most of the undefined terms will be explained in later sections.

A. The problem. Associated with any error-correcting code is a polynomial called its weight enumerator $W(x, y)$ (see Part II). For a certain class of codes this polynomial must satisfy two equations:

$$W\left(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}\right) = W(x, y), \tag{1}$$

$$W(x, iy) = W(x, y). \tag{2}$$

The problem we want to solve is to find all polynomials $W(x, y)$ satisfying (1) and (2).

B. Invariants. Equation (1) says that $W(x, y)$ is unchanged, or *invariant*, under the linear transformation

$$T_1:\quad \begin{array}{l}\text{replace } x \text{ by } (x+y)/\sqrt{2},\\ \text{replace } y \text{ by } (x-y)/\sqrt{2},\end{array}$$

or, in matrix notation,

$$T_1\text{: replace } \begin{pmatrix} x \\ y \end{pmatrix} \text{ by } \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Similarly Eq. (2) says that $W(x, y)$ is also invariant under the transformation

$$T_2:\quad \begin{array}{l}\text{replace } x \text{ by } x\\ \text{replace } y \text{ by } iy\end{array}$$

or

$$T_2\text{: replace } \begin{pmatrix} x \\ y \end{pmatrix} \text{ by } \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Of course $W(x, y)$ must therefore be invariant under any combination T_1^2, T_1T_2, $T_1T_2T_1, \ldots$ of these transformations. It is not difficult to show (as we shall see in Section A of Part III) that the matrices

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$$

when multiplied together in all possible ways produce a *group* $\mathfrak{G}_1$ containing 192 matrices.

So our problem now says: find the polynomials $W(x, y)$ which are invariant under all 192 matrices in the group $\mathfrak{G}_1$.

C. How many invariants? The first thing we want to know is how many invariants there are. This isn't too precise, because of course if f and g are invariants, so is any constant multiple cf and also $f+g$, $f-g$ and the product fg. Also it is enough to study the *homogeneous* invariants (in which all terms have the same degree).

So the right question to ask is: how many linearly independent, homogeneous invariants are there of each degree d? Let's call this number a_d.

A convenient way to handle the numbers $a_0, a_1, a_2, \ldots$ is by combining them into a power series or generating function

$$\Phi(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \cdots.$$

Conversely, if we know $\Phi(\lambda)$, the numbers a_d can be recovered from the power series expansion of $\Phi(\lambda)$.

At this point we invoke a beautiful theorem of T. Molien, published in 1897 ([**52**], [**14**, p. 301], [**51**, p. 259], [**11**, p. 110]):

THEOREM 1. *For any finite group $\mathfrak{G}$ of complex $m \times m$ matrices, $\Phi(\lambda)$ is given by*

$$\Phi(\lambda) = \frac{1}{|\mathfrak{G}|}\sum_{A\in\mathfrak{G}} \frac{1}{\det(I-\lambda A)}, \tag{3}$$

where $|\mathfrak{G}|$ is the number of matrices in $\mathfrak{G}$, det *stands for determinant, and I is a unit matrix. In words, $\Phi(\lambda)$ is the average, taken over all matrices A in the group, of the reciprocal of the polynomial* $\det(I-\lambda A)$.

We call $\Phi(\lambda)$ the *Molien series* of $\mathfrak{G}$. The proof of this theorem is given in Section D of Part III.

For our group $\mathfrak{G}_1$, from the matrices corresponding to I, T_1, $T_2, \ldots$ we get

$$\Phi(\lambda)=\frac{1}{192}\left\{\frac{1}{(1-\lambda)^2}+\frac{1}{1-\lambda^2}+\frac{1}{(1-\lambda)(1-i\lambda)}+\cdots\right\}. \tag{4}$$

There are shortcuts, but it is quite feasible to work out the 192 terms directly (many are the same) and add them. The result is a surprise: everything collapses to give

$$\Phi(\lambda)=\frac{1}{(1-\lambda^8)(1-\lambda^{24})}. \tag{5}$$

D. Interpreting $\Phi(\lambda)$. The very simple form of (5) is trying to tell us something. Expanding in powers of λ, we have

$$\begin{aligned}\Phi(\lambda)&=a_0+a_1\lambda+a_2\lambda^2+\cdots\\&=(1+\lambda^8+\lambda^{16}+\lambda^{24}+\cdots)(1+\lambda^{24}+\lambda^{48}+\cdots).\end{aligned} \tag{6}$$

We can deduce one fact immediately: a_d is zero unless d is a multiple of 8, i.e., the degree of a homogeneous invariant must be a multiple of 8. (This is already a useful theorem in coding theory.) But we can say more. The RHS of (6) is exactly what we would find if there were two "basic" invariants, of degrees 8 and 24, such that all invariants are formed from sums and products of them.

This is because two invariants, θ, of degree 8, and φ, of degree 24, would give rise to the following invariants.

(7)

degree d	invariants	number a_d
0	1	1
8	θ	1
16	θ^2	1
24	θ^3, φ	2
32	θ^4, $\theta\varphi$	2
40	θ^5, $\theta^2\varphi$	2
48	θ^6, $\theta^3\varphi$, φ^2	3
...	...	...

Provided all the products $\theta^i\varphi^j$ are linearly independent — which is the same thing as saying that θ and φ are algebraically independent — the numbers a_d in (7) are exactly the coefficients in

$$\begin{aligned}&1+\lambda^8+\lambda^{16}+2\lambda^{24}+2\lambda^{32}+2\lambda^{40}+3\lambda^{48}+\cdots\\&=(1+\lambda^8+\lambda^{16}+\lambda^{24}+\cdots)(1+\lambda^{24}+\lambda^{48}+\cdots)\\&=\frac{1}{(1-\lambda^8)(1-\lambda^{24})},\end{aligned} \tag{8}$$

which agrees with (5). So if we can find two algebraically independent invariants of degrees 8 and 24, we shall have solved our problem. The answer will be that any invariant of this group is a polynomial in θ and φ. We shall find θ and φ in the next section.

Notice how the exponents 8 and 24 in the denominator of (5) led us to guess the degrees of the basic invariants.

This behavior is typical, and is what makes the technique exciting to use. One starts with a group of matrices $\mathfrak{G}$, computes the complicated-looking sum shown in Eq. (3), and simplifies the result. Everything miraculously collapses, leaving a final expression resembling Eq. (5) (although not always quite so simple — the precise form of the final expression is given in Section E of Part III). This expression then tells you the degrees of the basic invariants to look for.

E. Finding the basic invariants. Finding the basic invariants is in general an easier problem than finding $\Phi(\lambda)$. There are two methods.

(a) *Finding invariants by averaging.* This method uses the following simple result (which is proved in Section B of Part II).

THEOREM 2. *If $f(\mathbf{x}) = f(x_1, \ldots, x_m)$ is any polynomial in m variables, and $\mathfrak{G}$ is a finite group of $m \times m$ matrices, then*

$$\bar{f}(\mathbf{x}) = \frac{1}{|\mathfrak{G}|} \sum_{A \in \mathfrak{G}} A \circ f(\mathbf{x})$$

is an invariant, where $A \circ f(\mathbf{x})$ denotes the polynomial obtained by applying the transformation A to the variables in f.

Of course $\bar{f}(\mathbf{x})$ may be zero. We shall give an example of the use of this theorem below. But in our present example the second method is easier to use.

(b) *The indirect method*, which is to use what we know about the problem to find invariants. In the present example we are studying self-dual codes with weights divisible by 4 (defined in Section B of Part II). Their weight enumerators satisfy Eqs. (1) and (2). There are two famous codes in this class, the extended Hamming code of length 8 and the extended Golay code of length 24 (codes C_6 and C_8 of Part II). The weight enumerators of these codes are respectively

$$\theta = x^8 + 14x^4y^4 + y^8 \tag{9}$$

and

$$\varphi' = x^{24} + 759x^{16}y^8 + 2576x^{12}y^{12} + 759x^8y^{16} + y^{24}. \tag{10}$$

Since the Hamming and Golay codes are self-dual, and have weights divisible by 4, these two polynomials must be invariant under Eqs. (1) and (2) and hence under the group $\mathfrak{G}_1$. So we have found the two basic invariants we were looking for. (It's not difficult to verify that they are algebraically independent.) Actually it is easier to work with

$$\varphi = \frac{\theta^3 - \varphi'}{42} = x^4y^4(x^4 - y^4)^4 \tag{10a}$$

rather than with φ'. So we have proved the following theorem, discovered by Gleason in 1970.

THEOREM 3a. *Any invariant of the group $\mathfrak{G}_1$ is a polynomial in θ (Eq. (9)) and φ (Eq. (10a)).*

This also gives us the solution to our original problem:

THEOREM 3b. *Any polynomial which satisfies Eqs. (1) and (2) is a polynomial in θ and φ.*

Finally, we have also obtained a very useful theorem about codes.

THEOREM 3c. (Gleason **[26]**) *The weight enumerator of any self-dual code, with weights divisible by 4, is a polynomial in θ and φ.*

Alternative proofs of this theorem are given in **[9]** and **[12]** (see also **[4]**). But the proof given here seems to be the most informative, and the easiest to understand and to generalize.

F. An application. To show how powerful Theorem 3c is, we shall use it to find the weight enumerator of the extended quadratic residue code of length 48 (the code C_9 of Part II).

All we need to know about this code is that it is self-dual, and that the weight of any nonzero codeword is a multiple of 4 and is at least 12 (i.e., this is a 5-error-correcting code). This implies that the weight enumerator of the code, which is a homogeneous polynomial of degree 48, has the form

$$W(x, y) = x^{48} + A_{12}x^{36}y^{12} + \cdots. \tag{11}$$

The coefficients of $x^{47}y$, $x^{46}y^2, \dots, x^{37}y^{11}$ are zero. Here A_{12} is the unknown number of codewords of weight 12. It is remarkable that, once we know Eq. (11), the weight enumerator is completely determined by Th. 3c. For Th. 3c says that $W(x, y)$ must be a polynomial in θ and φ. Since $W(x, y)$ is homogeneous of degree 48, θ is homogeneous of degree 8, and φ is homogeneous of degree 24, this polynomial must be a linear combination of θ^6, $\theta^3\varphi$, and φ^2.

Thus Th. 3c says that

$$W(x, y) = a_0\theta^6 + a_1\theta^3\varphi + a_2\varphi^2 \tag{12}$$

for some real numbers a_0, a_1, a_2. Expanding Eq. (12) we have

$$\begin{aligned} W(x, y) = a_0(x^{48} + 84x^{44}y^4 + 2946x^{40}y^8 + \cdots) \\ + a_1(x^{44}y^4 + 38x^{40}y^8 + \cdots) \\ + a_2(x^{40}y^8 - \cdots), \end{aligned} \tag{13}$$

and equating coefficients in Eqs. (11), (13) we get

$$a_0 = 1, \qquad a_1 = -84, \qquad a_2 = 246.$$

Therefore $W(x, y)$ is uniquely determined. When the values of a_0, a_1, a_2 are substituted in (12) it is found that

$$\begin{aligned} W(x, y) = x^{48} + 17296x^{36}y^{12} + 535095x^{32}y^{16} \\ + 3995376x^{28}y^{20} + 7681680x^{24}y^{24} + 3995376x^{20}y^{28} \\ + 535095x^{16}y^{32} + 17296x^{12}y^{36} + y^{48}. \end{aligned} \tag{14}$$

Direct calculation of this weight enumerator would require finding the weight of each of the $2^{24} \approx 1.7 \times 10^7$ codewords, a respectable job even for a computer.

Of course there is also a fair amount of algebra involved in the invariant theory method, although in the preceding example it can be done by hand. The reader may find it helpful if we give a second example, in which the algebra can be shown in full.

G. A very simple example. The weight enumerator of a self-dual code with symbols from the field of q elements must satisfy the equation

$$W\left(\frac{x + (q-1)y}{\sqrt{q}}, \frac{x - y}{\sqrt{q}}\right) = W(x, y). \tag{15}$$

Problem: Find all polynomials which satisfy Eq. (15).

The solution proceeds as before. Equation (15) says that $W(x, y)$ must be invariant under the transformation

$$T_3\text{: replace } \begin{pmatrix} x \\ y \end{pmatrix} \text{ by } A\begin{pmatrix} x \\ y \end{pmatrix},$$

where

$$A = \frac{1}{\sqrt{q}}\begin{pmatrix} 1 & q-1 \\ 1 & -1 \end{pmatrix}. \tag{16}$$

Now $A^2 = I$, so $W(x, y)$ must be invariant under the group $\mathfrak{G}_2$ consisting of the two matrices I and A.

To find how many invariants there are, we compute the Molien series $\Phi(\lambda)$ from Eq. (3). We find

$$\det(I-\lambda I)=(1-\lambda)^2,$$

$$\det(I-\lambda A)=\det\begin{pmatrix}1-\dfrac{\lambda}{\sqrt{q}} & -\dfrac{q-1}{\sqrt{q}}\lambda\\ -\dfrac{\lambda}{\sqrt{q}} & 1+\dfrac{\lambda}{\sqrt{q}}\end{pmatrix}=1-\lambda^2,$$

$$\Phi(\lambda)=\frac{1}{2}\left(\frac{1}{(1-\lambda)^2}+\frac{1}{1-\lambda^2}\right) \tag{17}$$

$$=\frac{1}{(1-\lambda)(1-\lambda^2)},$$

which is even simpler than Eq. (5). Equation (17) suggests that there might be two basic invariants, of degrees 1 and 2 (the exponents in the denominator). If algebraically independent invariants of degrees 1 and 2 can be found, say g and h, then Eq. (17) implies that any invariant of $\mathfrak{G}_2$ is a polynomial in g and h.

This time we shall use the first method (averaging) to find the basic invariants. Let us average x over the group—i.e., apply Theorem 2 with $f(x,y)=x$. The matrix I leaves x unchanged, of course, and the matrix A transforms x into $(1/\sqrt{q})(x+(q-1)y)$. Therefore the average,

$$\bar{f}(x,y)=\frac{1}{2}\left[x+\frac{1}{\sqrt{q}}\{x+(q-1)y\}\right]$$

$$=\frac{(\sqrt{q}+1)\{x+(\sqrt{q}-1)y\}}{2\sqrt{q}},$$

is an invariant. Of course any scalar multiple of $\bar{f}(x,y)$ is also an invariant, so we may divide by $(\sqrt{q}+1)/2\sqrt{q}$ and take

$$g=x+(\sqrt{q}-1)y \tag{18}$$

to be the basic invariant of degree 1. To get an invariant of degree 2 we average x^2 over the group, obtaining

$$\frac{1}{2}\left[x^2+\frac{1}{q}\{x+(q-1)y\}^2\right].$$

This can be cleaned up by subtracting $((q+1)/2q)g^2$ (which of course is an invariant), and dividing by a suitable constant. The result is

$$h=y(x-y),$$

the desired basic invariant of degree 2.

Finally, g and h must be shown to be algebraically independent: it must be shown that no sum of the form

$$\sum_{i,j} c_{ij}g^ih^j, \quad c_{ij} \text{ complex and not all zero,} \tag{19}$$

is identically zero when expanded in powers of x and y. This can be seen by looking at the leading terms. (The leading term of a polynomial is the first one to be written down when using the natural ordering illustrated in Eqs. (9), (14), (18).) Thus the leading term of g is x, the leading term of h is xy,

and the leading term of $g^i h^j$ is $x^{i+j} y^j$. Since distinct summands in Eq. (19) have distinct leading terms, (19) can only add to zero if all the c_{ij} are zero. Therefore g and h are algebraically independent. So we have proved:

THEOREM 4. *Any invariant of the group* $\mathfrak{G}_2$, *or equivalently any polynomial satisfying* (15), *or equivalently the weight enumerator of any self-dual code with symbols from* $GF(q)$, *is a polynomial in* $g = x + (\sqrt{q} - 1)y$ *and* $h = y(x - y)$.

At this point the coding theorist will cry "Stop!", and point out that a self-dual code must have even length and so every term in the weight enumerator must have even degree. But in Theorem 4 g has degree 1.

Thus we haven't made use of everything we know about the code. $W(x, y)$ must also be invariant under the transformation

$$\text{replace } \begin{pmatrix} x \\ y \end{pmatrix} \text{ by } B\begin{pmatrix} x \\ y \end{pmatrix},$$

where $B = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I$. This rules out terms of odd degree. So $W(x, y)$ is invariant under the group $\mathfrak{G}_3$ generated by A and B, which consists of

$$I, \ A, \ -I, \ -A.$$

The reader can easily work out that the new Molien series is

$$\begin{aligned} \Phi_{\mathfrak{G}_3}(\lambda) &= \frac{1}{2}\{\Phi_{\mathfrak{G}_2}(\lambda) + \Phi_{\mathfrak{G}_2}(-\lambda)\} \\ &= \frac{1}{2}\left\{\frac{1}{(1-\lambda)(1-\lambda^2)} + \frac{1}{(1+\lambda)(1-\lambda^2)}\right\} \\ &= \frac{1}{(1-\lambda^2)^2}. \end{aligned} \tag{20}$$

There are now two basic invariants, both of degree 2 (matching the exponents in the denominator of (20)), say g^2 and h, or the equivalent and slightly simpler pair $g^* = x^2 + (q-1)y^2$ and $h = y(x - y)$. Hence:

THEOREM 5. ([**41**]) *The weight enumerator of any self-dual code with symbols from the field of q elements is a polynomial in* g^* *and* h.

The preceding argument enables us to give a short proof of a recent result of Leontjev.

COROLLARY. (Leontjev [**35**]) *The weight enumerator* $W(x, y)$ *of a linear code over the field of q elements has the property that*

$$W(x, y)\, W\left(\frac{x + (q-1)y}{\sqrt{q}}, \frac{x - y}{\sqrt{q}}\right)$$

is a polynomial in g^* *and* h.

Proof. This product is clearly invariant under T_3 and $-I$, and so the result follows from the proof of Th. 5. Q.E.D.

H. The general plan of attack. As these examples have illustrated, there are two stages in using invariant theory to solve a problem.

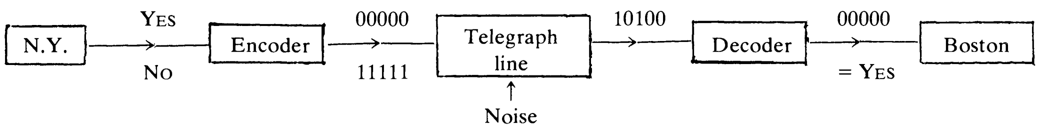
Stage I: Convert the assumptions about the problem (e.g., the code) into algebraic constraints on polynomials (e.g., weight enumerators).

Stage II: Use invariant theory to find all possible polynomials satisfying these constraints.

J. Arrangement of the paper. Part I has been devoted to explication by example. Part II gives the necessary background from coding theory, defines codes and weight enumerators, and explains why they are important. Part III is the main section of the paper and gives a brief account of invariant theory. Then Part IV gives further examples and illustrates more advanced techniques.

II. Background from Coding Theory

A. Definition of a code and examples. Imagine a noisy telegraph line from New York to Boston, which transmits 0's and 1's. Usually when a 0 is sent from New York it is received as a 0 in Boston, but occasionally a 0 is received as a 1. Similarly a 1 is occasionally received as a 0. The problem is to send a lot of important messages down this line, as quickly and as reliably as possible. The coding theorist's solution is to send only certain strings of 0's and 1's, called *codewords*. Here is a simple example: one of two messages will be sent, either YES or NO.



YES will be *encoded* into the codeword 00000, and NO into 11111. Suppose 10100 is received in Boston. The receiver argues that it is more likely that 00000 was sent (and two errors occurred) than that 11111 was sent (and three errors occurred), and therefore *decodes* 10100 as 00000 = YES. For in some sense 10100 is *closer* to 00000 than to 11111. To make this precise, define the *Hamming distance* (or simply the *distance*) between two vectors $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$ to be the number of places where $u_i \neq v_i$. This is denoted by $\operatorname{dist}(\mathbf{u}, \mathbf{v})$. E.g., $\operatorname{dist}(10100, 00000) = 2$, $\operatorname{dist}(10100, 11111) = 3$, and even $\operatorname{dist}(0122, 2001) = 4$ (the same definition applies to nonbinary vectors). It is easily checked that dist is a metric. Then the receiver should decode the received vector as the closest codeword, measured in Hamming distance.

Notice in this example two errors were corrected. This is possible because the codewords 00000 and 11111 are at distance 5 apart. In general, if d is the minimum Hamming distance between codewords, the code can correct $e = [\frac{1}{2}(d-1)]$ errors, where $[x]$ denotes the greatest integer not exceeding x. For if e or fewer errors occur, by the triangle inequality the received vector is still closer to the transmitted codeword than to any other. This motivates the

DEFINITION. An $[n, k, d]$ binary *code* consists of 2^k vectors $\mathbf{u} = (u_1, \ldots, u_n)$, $u_i \in F_2 = \{0, 1\}$, called *codewords*, such that

(i) the sum, taken componentwise and modulo 2 (without carries!), of any two codewords is again a codeword; and

(ii) any two codewords differ in at least d places.

Then n is called the *length*, k the *dimension*, and d the *minimum distance* of the code. In a good code n is small (for rapid transmission), k is large (for an efficient code), and d is large (to correct many errors). These are incompatible goals! For more about codes see for example [7], [8], [10], [36], [37], [42], [55].

Examples:

(1) $C_1 = \{00000, 11111\}$, the code of the example, is a $[5, 1, 5]$ code.

(2) More generally $C_2 = \{00\ldots0, 11\ldots1\}$ is an $[n, 1, n]$ *repetition* code.

(3) $C_3 = \{000, 011, 101, 110\}$ is a $[3, 2, 2]$ code. To verify this, note that each of the 4 codewords has 3 components. Also the sum of the codewords 011 and 101 (for example) is 110. Finally any two codewords differ in at least (in this case exactly) two coordinates.

(4) C_4, a $[7,3,4]$ code:

0000000
1110100
0111010
0011101
1001110
0100111
1010011
1101001

(5) C_5, a $[7,4,3]$ Hamming code, consists of the codewords of C_4 together with their complements:

0000000	1111111
1110100	0001011
0111010	1000101
0011101	1100010
1001110	0110001
0100111	1011000
1010011	0101100
1101001	0010110

(6) C_6, an $[8,4,4]$ extended Hamming code, formed by placing a 1 at the end of the 8 codewords on the left in C_5, and a 0 at the end of the 8 codewords on the right. (This is one of the codes mentioned in Part I, Section E.) The same technique can be applied to any $[n, k, d$ odd$]$ code. Placing a 1 at the end of the codewords containing an odd number of 1's and a 0 at the end of those with an even number of 1's we obtain an $[n+1, k, d+1]$ *extended* code.

For later use we mention a few important larger codes, although without giving any details.

(7) The two binary Golay codes, namely the $[23,12,7]$ code C_7 and the $[24,12,8]$ extended code C_8 (**[28]**, **[29]**).

(8) The codes C_5 and C_7 are both examples of *quadratic residue* codes (**[5]**, **[6]**). This is an infinite family of codes. Other quadratic residue codes are the $[31,16,7]$ code, the $[47,24,11]$ code, and the $[48,24,12]$ extended code C_9.

There are many situations in which it is better to use a nonbinary code. Let F_q denote the Galois field with q elements. E.g., $F_3 = \{0,1,2\}$ with addition, multiplication, division etc., performed modulo 3.

Definition. An $[n, k, d]$ code over F_q consists of q^k codewords $(u_1, \ldots, u_n)$, which have Hamming distance at least d apart and form a linear space. That is, the sum, performed componentwise in F_q, of any two codewords is again a codeword, and any scalar multiple $(cu_1, \ldots, cu_n)$, $c \in F_q$, of a codeword is again a codeword.

Examples:

(1) C_{10}, a $[4,2,2]$ code over F_3:

0000	1120	2102
0111	2210	2021
0222	1201	1012

(10) There is also a $[12,6,6]$ Golay code C_{11} over F_3—see **[28]**, **[29]**.

B. Dual code. Let $\mathscr{C}$ be an $[n, k, d]$ code over F_q. The *dual* (or *orthogonal*) code $\mathscr{C}^\perp$ consists of all vectors having zero dot product (in F_q) with every codeword of $\mathscr{C}$. Thus

$$\mathscr{C}^{\perp} = \left\{ (u_1, \ldots, u_n) \colon u \cdot v = \sum_{i=1}^{n} u_i v_i = 0 \right.$$
$$\left. \text{for all } v = (v_1, \ldots, v_n) \in \mathscr{C} \right\}.$$

Then it is easy to see that $\mathscr{C}^{\perp}$ is an $[n, n-k, d']$ code, for some positive integer d'.

Examples: The binary code $\{00, 11\}$ is its own dual! The dual of C_2 is the $[n, n-1, 2]$ code consisting of all codewords with an even number of 1's. The dual of C_3 is $\{000, 111\}$, and C_4 and C_5 are dual to each other. (It is an amusing and informative exercise to verify these facts.)

A *self-dual* code is one for which $\mathscr{C}^{\perp} = \mathscr{C}$. The examples $C_6, C_8, C_9, C_{10}, C_{11}$ are all self-dual. In a self-dual code k must be equal to $\frac{1}{2}n$, and so n must be even.

Self-dual codes are a particularly interesting class, for several reasons:

(1) It is known that there exist self-dual codes which are about as good as any code can be — more precisely, they meet the Gilbert–Varshamov bound **[44]**, **[60]**.

(2) A number of the best codes known are self-dual. For example, extended quadratic residue codes are self-dual whenever the length is of the form $n = 8m$ where $8m - 1$ is a prime.

(3) Since the codewords of the dual code $\mathscr{C}^{\perp}$ are parity check equations on $\mathscr{C}$, knowing that $\mathscr{C} = \mathscr{C}^{\perp}$ may simplify the decoding procedure (see for example **[27]**).

(4) The codewords in a self-dual code under certain conditions form t-designs, and many 5-designs have been constructed in this way (see **[3]**, **[4]**, **[56]**, **[57]**).

(5) There are connections between self-dual codes and sphere-packings, geometric lattices, large finite groups, and projective planes — see **[3]**, **[4]**, **[12]**, **[15]**–**[17]**, **[34]**, **[42]**, **[45]**, **[67]**.

Binary self-dual codes of length $n \leqq 24$ have been classified in **[58]**, **[61]**, **[62]**. See also **[24]**, **[50]**, **[59]**.

C. Weight enumerators. Examples (5)–(8) should have convinced the reader that codes are large, unwieldly objects. One way of handling a code and extracting some useful information from it is by means of its weight enumerator.

The *Hamming weight* (or simply the *weight*) of a vector $\mathbf{u} = (u_1, \ldots, u_n)$ is the number of nonzero u_i. This is denoted by $\mathrm{wt}(\mathbf{u})$. E.g., $\mathrm{wt}(1101000) = 3$, $\mathrm{wt}(1201) = 3$. Clearly $\mathrm{dist}(\mathbf{u}, \mathbf{v}) = \mathrm{wt}(\mathbf{u} - \mathbf{v})$. Since a code $\mathscr{C}$ is a linear space, for any codewords $\mathbf{u}, \mathbf{v}$, $\mathrm{dist}(\mathbf{u}, \mathbf{v}) = \mathrm{wt}(\mathbf{u} - \mathbf{v}) = \mathrm{wt}(\mathbf{w})$ for some $\mathbf{w} \in \mathscr{C}$. Therefore the minimum distance d between codewords is equal to the smallest weight of any nonzero codeword.

The weight enumerator of an $[n, k, d]$ code $\mathscr{C}$ is simply a polynomial which tells the number of codewords of each weight. If $\mathscr{C}$ contains A_i codewords of weight i, then the *weight enumerator* of $\mathscr{C}$ is defined to be

$$W_{\mathscr{C}}(x, y) = \sum_{i=0}^{n} A_i x^{n-i} y^i, \tag{21}$$

where x and y are indeterminates. Stated another way,

$$W_{\mathscr{C}}(x, y) = \sum_{\mathbf{u} \in \mathscr{C}} x^{n - \mathrm{wt}(\mathbf{u})} y^{\mathrm{wt}(\mathbf{u})}. \tag{22}$$

Equations (21) and (22) are complicated-looking definitions for a very simple idea, which some examples will make clear.

	Code $\mathscr{C}$	*Weight Enumerator* $W_{\mathscr{C}}(x, y)$
(23)	$\{00, 11\}$	$x^2 + y^2$
	C_2	$x^n + y^n$
	C_3	$x^3 + 3xy^2$

(24)	C_4	$x^7+7x^3y^4$
	C_5	$x^7+7x^4y^3+7x^3y^4+y^7$
(9)	C_6	$\theta=x^8+14x^4y^4+y^8$
(25)	C_7	x^4+8xy^3

For example, the weight enumerator of C_4 is $x^7+7x^3y^4$ because there is one codeword with 7 zeros (giving the term x^7) and 7 codewords with 3 zeros and 4 ones (giving the term $7x^3y^4$). The weight enumerators of C_8 and C_9 were given in Eqs. (10) and (14).

Notice that $W_{\mathscr{C}}(x,y)$ is a homogeneous polynomial of degree n. The weight enumerator immediately gives the minimum distance d of $\mathscr{C}$. For $\mathscr{C}$ always contains the zero codeword, giving the leading term x^n in $W_{\mathscr{C}}(x,y)$, and the next nonzero term is $A_d x^{n-d}y^d$. Thus

$$W_{\mathscr{C}}(x,y)=x^n+0x^{n-1}y+\cdots+0x^{n-d+1}y^{d-1}+A_d x^{n-d}y^d+\cdots.$$

An illustration was given in Eq. (11). $W_{\mathscr{C}}$ is also used to find the error probability of the code, and for other purposes—see [7] or [**42**].

D. MacWilliams theorem. For a large code it is in general a very tough problem to find the weight enumerator. One of the chief weapons available is the following remarkable theorem of F. J. MacWilliams, which states that the weight enumerator of the dual code $\mathscr{C}^\perp$ is uniquely determined by the weight enumerator of $\mathscr{C}$.

THEOREM 6. (MacWilliams [**40**]; see also [**42**], [**43**]). *If $\mathscr{C}$ is an $[n,k,d]$ code over F_q with dual code $\mathscr{C}^\perp$, then*

$$W_{\mathscr{C}^\perp}(x,y)=\frac{1}{q^k}W_{\mathscr{C}}(x+(q-1)y,x-y). \tag{26}$$

We shall just prove the binary version, when $q=2$. This states that

$$W_{\mathscr{C}^\perp}(x,y)=\frac{1}{2^k}W_{\mathscr{C}}(x+y,x-y), \tag{27}$$

or equivalently

$$\sum_{\mathbf{u}\in\mathscr{C}^\perp}x^{n-\mathrm{wt}(\mathbf{u})}y^{\mathrm{wt}(\mathbf{u})}=\frac{1}{2^k}\sum_{\mathbf{u}\in\mathscr{C}}(x+y)^{n-\mathrm{wt}(\mathbf{u})}(x-y)^{\mathrm{wt}(\mathbf{u})}. \tag{28}$$

The proof depends on the following lemma, which is in fact a version of the Poisson summation formula [**23**, p. 220]. Let F^n denote the set of all binary vectors of length n.

LEMMA 7. ([**37**]). *Let f be any mapping defined on F^n. We must be able to add and subtract the values $f(\mathbf{u})$, but otherwise f can be arbitrary. The* **Hadamard transform** *of f, $\hat{f}$, is defined by*

$$\hat{f}(\mathbf{u})=\sum_{\mathbf{v}\in F^n}(-1)^{\mathbf{u}\cdot\mathbf{v}}f(\mathbf{v}),\qquad \mathbf{u}\in F^n. \tag{29}$$

Then if $\mathscr{C}$ is an $[n,k,d]$ binary code

$$\sum_{\mathbf{u}\in\mathscr{C}^\perp}f(\mathbf{u})=\frac{1}{2^k}\sum_{\mathbf{u}\in\mathscr{C}}\hat{f}(\mathbf{u}). \tag{30}$$

Proof.

$$\sum_{\mathbf{u}\in\mathscr{C}}\hat{f}(\mathbf{u})=\sum_{\mathbf{u}\in\mathscr{C}}\sum_{\mathbf{v}\in F^n}(-1)^{\mathbf{u}\cdot\mathbf{v}}f(\mathbf{v})=\sum_{\mathbf{v}\in F^n}f(\mathbf{v})\sum_{\mathbf{u}\in\mathscr{C}}(-1)^{\mathbf{u}\cdot\mathbf{v}}.$$

Now if $\mathbf{v}\in\mathscr{C}^\perp$, $\mathbf{u}\cdot\mathbf{v}$ is always zero, and the inner sum is 2^k. But if $\mathbf{v}\notin\mathscr{C}^\perp$ then a moment's thought

shows that $\mathbf{u}\cdot\mathbf{v}$ is 0 and 1 equally often, and the inner sum is 0. Therefore

$$\sum_{\mathbf{u}\in\mathscr{C}} \hat{f}(\mathbf{u}) = 2^k \sum_{\mathbf{v}\in\mathscr{C}^\perp} f(\mathbf{v}). \qquad \text{Q.E.D.}$$

Proof of Theorem 6. We apply the lemma with

$$\begin{aligned} f(\mathbf{u}) &= x^{n-\mathrm{wt}(\mathbf{u})}y^{\mathrm{wt}(\mathbf{u})}, \\ \hat{f}(\mathbf{u}) &= \sum_{\mathbf{v}\in F^n} (-1)^{\mathbf{u}\cdot\mathbf{v}} x^{n-\mathrm{wt}(\mathbf{v})}y^{\mathrm{wt}(\mathbf{v})}. \end{aligned} \tag{31}$$

Let $\mathbf{u} = (u_1\cdots u_n)$, $\mathbf{v} = (v_1\cdots v_n)$. Then

$$\begin{aligned} \hat{f}(\mathbf{u}) &= \sum_{\mathbf{v}\in F^n} (-1)^{u_1v_1+\cdots+u_nv_n} \prod_{i=1}^{n} x^{1-v_i}y^{v_i} \\ &= \sum_{v_1=0}^{1}\cdots\sum_{v_n=0}^{1}\prod_{i=1}^{n}(-1)^{u_iv_i}x^{1-v_i}y^{v_i}. \end{aligned} \tag{32}$$

Just as

$$a_0b_0c_0 + a_0b_0c_1 + a_0b_1c_0 + a_0b_1c_1 + a_1b_0c_0 + a_1b_0c_1 + a_1b_1c_0 + a_1b_1c_1$$

is equal to $(a_0+a_1)(b_0+b_1)(c_0+c_1)$, so (32) is equal to

$$\prod_{i=1}^{n}\sum_{w=0}^{1}(-1)^{u_iw}x^{1-w}y^{w}.$$

If $u_i = 0$, the inner sum is $x+y$. If $u_i = 1$, it is $x-y$. Thus

$$\hat{f}(\mathbf{u}) = (x+y)^{n-\mathrm{wt}(\mathbf{u})}(x-y)^{\mathrm{wt}(\mathbf{u})}. \tag{33}$$

Then Eq. (30) reads

$$\sum_{\mathbf{u}\in\mathscr{C}^\perp} x^{n-\mathrm{wt}(\mathbf{u})}y^{\mathrm{wt}(\mathbf{u})} = \frac{1}{2^k}\sum_{\mathbf{u}\in\mathscr{C}}(x+y)^{n-\mathrm{wt}(\mathbf{u})}(x-y)^{\mathrm{wt}(\mathbf{u})},$$

which is Eq. (28). Q.E.D.

Examples: Let us illustrate Th. 6 by applying it to the code $\mathscr{C} = C_3$ with weight enumerator $W_{C_3}(x,y) = x^3 + 3xy^2$. Then

$$\tfrac{1}{4}W_{C_3}(x+y, x-y) = \tfrac{1}{4}\{(x+y)^3 + 3(x+y)(x-y)^2\} = x^3 + y^3,$$

which is indeed the weight enumerator of the dual code $C_3^\perp = \{000, 111\}$.

Of course if $\mathscr{C}$ is a self-dual code, both sides of Eqs. (26) and (27) must be the same. For example, if $\mathscr{C} = \{00, 11\}$, $W_{\mathscr{C}}(x,y) = x^2 + y^2$, and

$$\begin{aligned} \tfrac{1}{2}W_{\mathscr{C}}(x+y, x-y) &= \tfrac{1}{2}\{(x+y)^2 + (x-y)^2\} \\ &= x^2 + y^2 = W_{\mathscr{C}}(x,y) \end{aligned}$$

which is correct since $\mathscr{C}$ is self-dual. This is an illustration of

COROLLARY 8. *If $\mathscr{C}$ is an $[n, \frac{1}{2}n, k]$ self-dual code over F_q, then*

$$W_{\mathscr{C}}(x,y) = \frac{1}{q^{n/2}}W_{\mathscr{C}}(x+(q-1)y, x-y).$$

Since $W_{\mathscr{C}}(x,y)$ is a homogeneous polynomial of degree n, we can write this as

(34) $$W_{\mathscr{C}}(x, y) = W_{\mathscr{C}}\left(\frac{x+(q-1)y}{\sqrt{q}}, \frac{x-y}{\sqrt{q}}\right).$$

We have already used this fact in Part I, Section G, where we pointed out that (34) implies that $W_{\mathscr{C}}(x, y)$ is *invariant* under a certain linear transformation. This brings us to the main part of the paper.

III. Invariant Theory

A. Groups of matrices. Throughout this paper the letters $\mathfrak{G}, \mathfrak{H}, \ldots$ denote finite groups of complex $m \times m$ matrices. We remind the reader that the statement "$\mathfrak{G}$ is a group of matrices" means that $\mathfrak{G}$ is set of invertible matrices with the following properties: if A and B are in $\mathfrak{G}$ so is the product AB; the unit matrix I is in $\mathfrak{G}$; and the inverse A^{-1} of every $A \in \mathfrak{G}$ is also in $\mathfrak{G}$. The number of matrices in $\mathfrak{G}$ is called its *order* and will be denoted by g.

Given a collection $A_1, \ldots, A_r$ of $m \times m$ matrices we can form a group $\mathfrak{G}$ from them by multiplying them together in all possible ways. Thus $\mathfrak{G}$ contains the matrices I, A_1, $A_2, \ldots$, $A_1A_2, \ldots$, $A_2A_1^{-1}A_2^{-1}A_3, \ldots$. We say that $\mathfrak{G}$ is *generated* by $A_1, \ldots, A_r$. Of course $\mathfrak{G}$ may be infinite, in which case the theory of invariants described here doesn't directly apply. (But see **[20]**, **[63]**, **[74]**.)

Example: Let us show that the group $\mathfrak{G}_1$ generated by the matrices

$$M = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix} \text{ and } J = \begin{pmatrix}1 & 0\\ 0 & i\end{pmatrix}$$

that was encountered in Section B of Part I does indeed have order 192. The key is to discover (by randomly multiplying matrices together) that $\mathfrak{G}_1$ contains

$$J^2 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \qquad E = (MJ)^3 = \frac{1+i}{\sqrt{2}}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},$$

$$E^2 = -i\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \qquad R = MJ^2M = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}.$$

Set $\eta = (1+i)/\sqrt{2} = \cos 45° + i \sin 45°$. Then $\mathfrak{G}_1$ contains the 16 matrices

$$\alpha\begin{pmatrix}1 & 0\\ 0 & \pm 1\end{pmatrix}, \qquad \alpha\begin{pmatrix}0 & 1\\ \pm 1 & 0\end{pmatrix}, \qquad \alpha \in \{1, i, -1, -i\},$$

which form a subgroup $\mathfrak{H}_1$. From this it is easy to see that $\mathfrak{G}_1$ consists of the union of $\mathfrak{H}_1$ and 11 *cosets* $a_k\mathfrak{H}_1 = \{a_kA : A \in \mathfrak{H}_1\}$. Thus

(35) $$\mathfrak{G}_1 = \bigcup_{k=1}^{12} a_k\mathfrak{H}_1,$$

where $a_1, \ldots, a_6$ are respectively

$$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & i\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ i & -i\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1 & i\\ 1 & -i\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}1 & i\\ i & 1\end{pmatrix},$$

and $a_7 = \eta a_1, \ldots, a_{12} = \eta a_6$. From this it is possible to obtain a list of all 192 matrices in $\mathfrak{G}_1$—they are the matrices

(36) $$\eta^\nu\begin{pmatrix}1 & 0\\ 0 & \alpha\end{pmatrix}, \quad \eta^\nu\begin{pmatrix}0 & 1\\ \alpha & 0\end{pmatrix}, \quad \eta^\nu\frac{1}{\sqrt{2}}\begin{pmatrix}1 & \beta\\ \alpha & -\alpha\beta\end{pmatrix},$$

for $0 \leqq \nu \leqq 7$ and $\alpha, \beta \in \{1, i, -1, -i\}$.

As a check, one verifies that every matrix in (36) can be written as a product of M's and J's; that

the product of two matrices in (36) is again in (36); and that the inverse of every matrix in (36) is in (36). Therefore (36) *is* a group, and is the group generated by M and J. Thus $\mathfrak{G}_1$ is indeed equal to (36).

We have gone into this example in some detail to emphasize that it is important to begin by understanding the group thoroughly. (For an alternative way of studying $\mathfrak{G}_1$ see [**12**; pp. 160–161].)

B. Invariants. To quote Hermann Weyl [**73**], "the theory of invariants came into existence about the middle of the nineteenth century somewhat like Minerva: a grown-up virgin, mailed in the shining armor of algebra, she sprang forth from Cayley's Jovian head." Invariant theory became one of the main branches of nineteenth century mathematics, but dropped out of fashion after Hilbert's work: see [**25**], [**64**]. Recently, however, there has been a resurgence of interest, with applications in algebraic geometry [**20**], [**53**], physics (see for example [**1**] and the references given there), combinatorics [**22**], [**65**], and coding theory [**41**], [**48**], [**49**].

There are several different kinds of invariants, but in this paper an invariant is defined as follows.

Let $\mathfrak{G}$ be a group of g $m \times m$ complex matrices $A_1, \ldots, A_g$, where the (i, k)th entry of A_α is $a_{ik}^{(\alpha)}$. In other words, $\mathfrak{G}$ is a group of linear transformations on the variables $x_1, \ldots, x_m$, consisting of the transformations

$$T^{(\alpha)}\text{: replace } x_i \text{ by } x_i^{(\alpha)} = \sum_{k=1}^{m} a_{ik}^{(\alpha)} x_k, \quad i = 1, \ldots, m \tag{37}$$

for $\alpha = 1, 2, \ldots, g$. It is worthwhile giving a careful description of how a polynomial $f(\mathbf{x}) = f(x_1, \ldots, x_m)$ is transformed by a matrix A_α in $\mathfrak{G}$. The transformed polynomial is

$$A_\alpha \circ f(\mathbf{x}) = f(x_1^{(\alpha)}, \ldots, x_m^{(\alpha)}),$$

where each $x_i^{(\alpha)}$ is replaced by $\sum_{k=1}^{m} a_{ik}^{(\alpha)} x_k$. Another way of describing this is to think of $\mathbf{x} = (x_1, \ldots, x_m)^T$ as a column vector (where the T denotes transpose). Then $f(\mathbf{x})$ is transformed into

$$A_\alpha \circ f(\mathbf{x}) = f(A_\alpha \mathbf{x}), \tag{38}$$

where $A_\alpha \mathbf{x}$ is the usual product of a matrix and a vector. One can check that

$$B \circ (A \circ f(\mathbf{x})) = (AB) \circ f(\mathbf{x}) = f(AB\mathbf{x}). \tag{39}$$

For example, $A = \left(\begin{smallmatrix} 1 & 2 \\ 0 & -1 \end{smallmatrix}\right)$ transforms $x_1^2 + x_2$ into $(x_1 + 2x_2)^2 - x_2$.

Definition. An *invariant* of $\mathfrak{G}$ is a polynomial $f(\mathbf{x})$ which is unchanged by every linear transformation in $\mathfrak{G}$. In other words, $f(\mathbf{x})$ is an invariant of $\mathfrak{G}$ if

$$A_\alpha \circ f(\mathbf{x}) = f(A_\alpha \mathbf{x}) = f(\mathbf{x})$$

for all $\alpha = 1, \ldots, g$.

Example: Let

$$\mathfrak{G}_4 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\},$$

a group of order $g = 2$. Then x^2, xy and y^2 are homogeneous invariants of degree 2.

Even if $f(\mathbf{x})$ isn't an invariant, its average over the group is, as was mentioned in Section E of Part I.

Theorem 2. *Let $f(\mathbf{x})$ be any polynomial. Then*

$$\bar{f}(\mathbf{x}) = \frac{1}{g} \sum_{\alpha=1}^{g} A_\alpha \circ f(\mathbf{x}) \tag{40}$$

is an invariant of $\mathfrak{G}$.

Proof. Any $A_\beta \in \mathfrak{G}$ transforms the right-hand side of (40) into

$$\frac{1}{g}\sum_{\alpha=1}^{g}(A_\alpha A_\beta)\circ f(\mathbf{x}), \quad \text{by (39)}. \tag{41}$$

It is an easy exercise in group theory to see that as A_α runs through $\mathfrak{G}$, so does $A_\alpha A_\beta$, if A_β is fixed. Therefore (41) is equal to

$$\frac{1}{g}\sum_{\gamma=1}^{g} A_\gamma \circ f(\mathbf{x})$$

which is $\bar{f}(\mathbf{x})$. Therefore $\bar{f}(\mathbf{x})$ is an invariant. Q.E.D.

More generally, any symmetric function of the g polynomials $A_1 \circ f(\mathbf{x}), \ldots, A_g \circ f(\mathbf{x})$ is an invariant of $\mathfrak{G}$.

Clearly if $f(\mathbf{x})$ and $h(\mathbf{x})$ are invariants of $\mathfrak{G}$, so are $f(\mathbf{x}) + h(\mathbf{x})$, $f(\mathbf{x})h(\mathbf{x})$, and $cf(\mathbf{x})$ (c complex). This is equivalent to saying that the set of invariants of $\mathfrak{G}$, which we denote by $\mathcal{I}(\mathfrak{G})$, forms a *ring.*

One of the main problems of invariant theory is to describe $\mathcal{I}(\mathfrak{G})$. Since the transformations in $\mathfrak{G}$ don't change the degree of a polynomial, it is enough to describe the homogeneous invariants (for any invariant is a sum of homogeneous invariants).

C. Basic invariants. Our goal is to find a "basis" for the invariants of $\mathfrak{G}$, that is, a set of basic invariants such that any invariant can be expressed in terms of this set. There are three different types of bases one might look for.

DEFINITION. Polynomials $f_1(\mathbf{x}), \ldots, f_r(\mathbf{x})$ are called *algebraically dependent* if there is a polynomial p in r variables with complex coefficients, not all zero, such that $p(f_1(\mathbf{x}), \ldots, f_r(\mathbf{x})) \equiv 0$. Otherwise $f_1(\mathbf{x}), \ldots, f_r(\mathbf{x})$ are called *algebraically independent.* A fundamental result from algebra is:

THEOREM 9 [**32**, p. 154]. *Any $m+1$ polynomials in m variables are algebraically dependent.*

The first type of basis we might look for is a set of m algebraically independent invariants $f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})$. Such a set is indeed a "basis", for by Th. 9 any invariant is algebraically dependent on $f_1, \ldots, f_m$ and so is a root of a polynomial equation in $f_1, \ldots, f_m$. The following theorem guarantees the existence of such a basis.

THEOREM 10 [**14**, p. 357]. *There always exist m algebraically independent invariants of $\mathfrak{G}$.*

Proof. Consider the polynomial $\Pi_{\alpha=1}^{g}(t - A_\alpha \circ x_1)$ in the variables $t, x_1, \ldots, x_m$. Since one of the A_α is the identity matrix, $t = x_1$ is a zero of this polynomial. When the polynomial is expanded in powers of t, the coefficients are invariants by the remark immediately following the proof of Th. 2. Therefore x_1 is an algebraic function of invariants. Similarly each of $x_2, \ldots, x_m$ is an algebraic function of invariants. Now if the number of algebraically independent invariants were m' ($< m$), the m independent variables $x_1, \ldots, x_m$ would be algebraic functions of the m' invariants, a contradiction. Therefore the number of algebraically independent invariants is at least m. By Th. 9 this number cannot be greater than m. Q.E.D.

Example: For the preceding group $\mathfrak{G}_4$, we may take $f_1 = (x+y)^2$ and $f_2 = (x-y)^2$ as the algebraically independent invariants. Then any invariant is a root of a polynomial equation in f_1 and f_2. For example,

$$x^2 = \tfrac{1}{4}(\sqrt{f_1} + \sqrt{f_2})^2, \quad xy = \tfrac{1}{4}(f_1 - f_2),$$

and so on.

The second type of basis, whose existence is guaranteed by the next theorem, is easier to work with.

THEOREM 11 [**14**, p. 359]. *There always exist* $m+1$ *invariants* $f_1, \ldots, f_{m+1}$ *of* $\mathfrak{G}$ *such that any invariant of* $\mathfrak{G}$ *is a rational function in* $f_1, \ldots, f_{m+1}$, *i.e., is the ratio of two polynomials in* $f_1, \ldots, f_{m+1}$.

Example: For $\mathfrak{G}_4$, x^2, xy and y^2 form such a basis.

However, by far the most convenient description of the invariants is a set $f_1, \ldots, f_l$ of invariants with the property that any invariant is a *polynomial* in $f_1, \ldots, f_l$. Then $f_1, \ldots, f_l$ is called a *polynomial basis* (or an *integrity basis*) for the invariants of $\mathfrak{G}$. Of course if $l > m$, then by Th. 9 there will be polynomial equations, called *syzygies*, relating $f_1, \ldots, f_l$.

For example, $f_1 = x^2$, $f_2 = xy$, $f_3 = y^2$ form a polynomial basis for the invariants of $\mathfrak{G}_4$. The syzygy relating them is $f_1 f_3 - f_2^2 = 0$. The existence of a polynomial basis, and a method of finding it, is given by the next theorem.

THEOREM 12 (Noether [**54**]; see also Weyl [**74**, p. 275]). *The ring of invariants of a finite group* $\mathfrak{G}$ *of complex* $m \times m$ *matrices has a polynomial basis consisting of not more than* $\binom{m+g}{m}$ *invariants, of degree not exceeding* g, *where* g *is the order of* $\mathfrak{G}$. *Furthermore, this basis may be obtained by taking the average over* $\mathfrak{G}$ *of all monomials*

$$x_1^{b_1} x_2^{b_2} \cdots x_m^{b_m}$$

of total degree $\Sigma\, b_i$ *not exceeding* g.

Proof. Let the group $\mathfrak{G}$ consist of the transformations (37). Suppose

$$f(x_1, \ldots, x_m) = \sum_{\mathbf{e}} c_{\mathbf{e}} x_1^{e_1} \cdots x_m^{e_m}, \qquad c_{\mathbf{e}} \text{ complex},$$

is any invariant of $\mathfrak{G}$. (The sum extends over all $\mathbf{e} = e_1, \ldots, e_m$ for which there is a nonzero term $x_1^{e_1} \cdots x_m^{e_m}$ in $f(x_1, \ldots, x_m)$.) Since $f(x_1, \ldots, x_m)$ is an invariant, it is unchanged when we average it over the group, so

$$\begin{aligned} f(x_1, \ldots, x_m) &= \frac{1}{g}\{f(x_1^{(1)}, \ldots, x_m^{(1)}) + \cdots + f(x_1^{(g)}, \ldots, x_m^{(g)})\} \\ &= \frac{1}{g} \sum_{\mathbf{e}} c_{\mathbf{e}} \{(x_1^{(1)})^{e_1} \cdots (x_m^{(1)})^{e_m} + \cdots \\ &\qquad \cdots + (x_1^{(g)})^{e_1} \cdots (x_m^{(g)})^{e_m}\} = \frac{1}{g} \sum_{\mathbf{e}} c_{\mathbf{e}} J_{\mathbf{e}} \qquad \text{(say)}. \end{aligned}$$

Every invariant is therefore a linear combination of the (infinitely many) special invariants

$$J_{\mathbf{e}} = \sum_{\alpha=1}^{g} (x_1^{(\alpha)})^{e_1} \cdots (x_m^{(\alpha)})^{e_m}.$$

Now $J_{\mathbf{e}}$ is (apart from a constant factor) the coefficient of $u_1^{e_1} \cdots u_m^{e_m}$ in

$$S_e = \sum_{\alpha=1}^{g} (u_1 x_1^{(\alpha)} + \cdots + u_m x_m^{(\alpha)})^e,$$

where $e = e_1 + \cdots + e_m$. In other words, the S_e are the *power sums* of the g quantities

$$u_1 x_1^{(1)} + \cdots + u_m x_m^{(1)}, \ldots, u_1 x_1^{(g)} + \cdots + u_m x_m^{(g)}.$$

It is well known [**33**] that any power sum S_e, $e = 1, 2, \ldots$, can be written as a polynomial with rational coefficients in the first g power sums

$$S_1, S_2, \ldots, S_g.$$

Therefore any $J_{\mathbf{e}}$ for $e = \Sigma_{i=1}^m e_i > g$ (which is a coefficient of S_e) can be written as a polynomial in the

special invariants

$$J_{\mathbf{e}} \text{ with } e_1 + \cdots + e_m \leqq g$$

(which are the coefficients of $S_1, \ldots, S_g$). Thus any invariant can be written as a polynomial in the $J_{\mathbf{e}}$ with $\Sigma_{i=1}^{m} e_i \leqq g$. The number of such $J_{\mathbf{e}}$ is the number of $e_1 e_2 \cdots e_m$ with $e_i \geqq 0$ and $e_1 + \cdots + e_m \leqq g$, which is $\binom{m+g}{m}$. Finally $\deg J_{\mathbf{e}} \leqq g$, and $J_{\mathbf{e}}$ is obtained by averaging $x_1^{e_1} \cdots x_m^{e_m}$ over the group. Q.E.D.

D. Molien's theorem. Since we know from Th. 12 that a polynomial basis always exists, we can go ahead with confidence and try to find it, using the methods described in Section E of Part I. To discover when a basis has been found, we use Molien's theorem (Theorem 1 above). This states that if a_d is the number of linearly independent homogeneous invariants of $\mathfrak{G}$ with degree d, and

$$\Phi_{\mathfrak{G}}(\lambda) = \sum_{d=0}^{\infty} a_d \lambda^d,$$

then

$$\Phi_{\mathfrak{G}}(\lambda) = \frac{1}{g} \sum_{\alpha=1}^{g} \frac{1}{\det(I - \lambda A_\alpha)}. \tag{42}$$

The proof depends on the following theorem.

THEOREM 13 [**51**, p. 258; **66**, p. 29]. *The number of linearly independent invariants of* $\mathfrak{G}$ *of degree* 1 *is*

$$a_1 = \frac{1}{g} \sum_{\alpha=1}^{g} \text{trace } (A_\alpha).$$

Proof. Let $S = (1/g)\Sigma_{\alpha=1}^{g} A_\alpha$. Changing the variables on which $\mathfrak{G}$ acts from $x_1, \ldots, x_m$ to $y_1, \ldots, y_m$, where $(y_1, \ldots, y_m) = (x_1, \ldots, x_m)T''$, changes S to $S' = TST^{-1}$. We may choose T so that S' is diagonal (see [**14**, p. 252]). Now $S^2 = S$, $(S')^2 = S'$, hence the diagonal entries of S' are 0 or 1. So with a change of variables we may assume

$$S = \begin{bmatrix} 1 & & & & & & \\ & \cdot & & & & 0 & \\ & & \cdot & & & & \\ & & & 1 & & & \\ & & & & 0 & & \\ & & & & & \cdot & \\ & 0 & & & & & \cdot \\ & & & & & & 0 \end{bmatrix},$$

with say r 1's on the diagonal. Thus $S \circ y_i = y_i$ if $1 \leqq i \leqq r$, $S \circ y_i = 0$ if $r + 1 \leqq i \leqq m$.

Any linear invariant of $\mathfrak{G}$ is certainly fixed by S, so $a_1 \leqq r$. On the other hand, by Th. 2, $S \circ y_i = (1/g)\Sigma_{\alpha=1}^{g} A_\alpha \circ y_i$ is an invariant of $\mathfrak{G}$ for any i, and so $a_1 \geqq r$. Q.E.D.

Before proving Th. 1 let us introduce some more notation. Equation (37) describes how A_α transforms the variables $x_1, \ldots, x_m$. The *d-th induced matrix*, denoted by $A_\alpha^{[d]}$, describes how A_α transforms the products of the x_i taken d at a time, namely $x_1^d, x_2^d, \ldots, x_1^{d-1}x_2, \ldots$ ([**2**], [**39**, p. 122]). E.g., $A_\alpha = \binom{a\ b}{c\ d}$ transforms x_1^2, x_1x_2 and x_2^2 into

$$x^2x_1^2 + 2abx_1x_2 + b^2x_2^2,$$
$$xcx_1^2 + (ad + bc)x_1x_2 + bdx_2^2,$$
$$c^2x_1^2 + 2cdx_1x_2 + d^2x_2^2$$

respectively. Thus the 2nd induced matrix is

$$A_\alpha^{[2]} = \begin{bmatrix} a^2 & 2ab & b^2 \\ ac & ad+bc & bd \\ c^2 & 2cd & d^2 \end{bmatrix}.$$

Proof of Theorem 1: To prove Eq. (42), note that a_d is equal to the number of linearly independent invariants of degree 1 of $\mathfrak{G}^{[d]} = \{A_\alpha^{[d]}: \alpha = 1, \ldots, g\}$. By Th. 13,

$$a_d = \frac{1}{g} \sum_{d=1}^{g} \text{trace } A_\alpha^{[d]}.$$

Therefore to prove Th. 1 it is enough to show that the trace of $A_\alpha^{[d]}$ is equal to the coefficient of λ^d in

$$\frac{1}{\det(I - \lambda A_\alpha)} = \frac{1}{(1-\lambda w_1)\cdots(1-\lambda w_m)}, \tag{43}$$

where $w_1, \ldots, w_m$ are the eigenvalues of A_α. By a suitable change of variables we can make

$$A_\alpha = \begin{bmatrix} w_1 & & & 0 \\ & \cdot & & \\ & & \cdot & \\ 0 & & & w_m \end{bmatrix}, \qquad A_\alpha^{[d]} = \begin{bmatrix} w_1^d & & & & & & 0 \\ & w_2^d & & & & & \\ & & \cdot & & & & \\ & & & \cdot & & & \\ & & & & \cdot & & \\ & & & & & w_1^{d-1} w_2 & \\ & & & & & & \cdot \\ 0 & & & & & & & \cdot \\ & & & & & & & & \cdot \end{bmatrix},$$

and trace $A_\alpha^{[d]}$ = sum of the products of $w_1, \ldots, w_m$ taken d at a time. But this is exactly the coefficient of λ^d in the expansion of (43). Q.E.D.

It is worth remarking that the Molien series does not determine the group. For example there are two groups of 2×2 matrices with order 8 having

$$\Phi(\lambda) = \frac{1}{(1-\lambda^2)(1-\lambda^4)}$$

(namely the dihedral group $\mathfrak{D}_4$ and the abelian group $\mathfrak{Z}_2 \times \mathfrak{Z}_4$). In fact there exist abstract groups $\mathfrak{A}$ and $\mathfrak{B}$ whose matrix representations can be paired in such a way that every representation of $\mathfrak{A}$ has the same Molien series as the corresponding representation of $\mathfrak{B}$ (**[18]**).

E. A standard form for the basic invariants. The following notation is very useful in describing the ring $\mathcal{J}(\mathfrak{G})$ of invariants of a group $\mathfrak{G}$. The complex numbers are denoted by $\mathbb{C}$, and if $p(\mathbf{x}), q(\mathbf{x}), \ldots$ are polynomials $\mathbb{C}[p(\mathbf{x}), q(\mathbf{x}), \ldots]$ denotes the set of all polynomials in $p(\mathbf{x})$, $q(\mathbf{x})$ with complex coefficients. For example Th. 3a just says that $\mathcal{J}(\mathfrak{G}_1) = \mathbb{C}[\theta, \varphi]$.

Also $\oplus$ will denote the usual direct sum operation. For example, a statement like $\mathcal{J}(\mathfrak{G}) = R \oplus S$ means that every invariant of $\mathfrak{G}$ can be written uniquely in the form $r + s$ where $r \in R$, $s \in S$. (Theorem 15 below illustrates this.)

Using this notation we can now specify the most convenient form of polynomial basis for $\mathcal{J}(\mathfrak{G})$.

DEFINITION. A *good polynomial basis* for $\mathcal{J}(\mathfrak{G})$ consists of homogeneous invariants $f_1, \ldots, f_l$ ($l \geqq m$) where $f_1, \ldots, f_m$ are algebraically independent and

$$\mathcal{J}(\mathfrak{G}) = \mathbb{C}[f_1, \ldots, f_m] \text{ if } l = m, \tag{44a}$$

or, if $l > m$,

(44b) $$\mathcal{I}(\mathfrak{G}) = \mathbb{C}[f_1,\ldots,f_m] \oplus f_{m+1}\mathbb{C}[f_1,\ldots,f_m] \oplus \cdots \oplus f_l\,\mathbb{C}[f_1,\ldots,f_m].$$

In words, this says that any invariant of $\mathfrak{G}$ can be written as a polynomial in $f_1,\ldots,f_m$ (if $l = m$), or as such a polynomial plus f_{m+1} times another such polynomial plus ... (if $l > m$). Speaking loosely, this says that, to describe an arbitrary invariant, $f_1,\ldots,f_m$ are "free" invariants and can be used as often as needed, while $f_{m+1},\ldots,f_l$ are "transient" invariants and can each be used at most once.

For a good polynomial basis $f_1,\ldots,f_l$ we can say exactly what the syzygies are. If $l = m$ there are no syzygies. If $l > m$ there are $(l-m)^2$ syzygies expressing the products f_if_j ($i \geqq m$, $j \geqq m$) in terms of $f_1,\ldots,f_l$.

It is important to note that the Molien series can be written down by inspection from the degrees of a good polynomial basis. Let $d_1 = \deg f_1,\ldots,d_l = \deg f_l$. Then

(45a) $$\Phi_{\mathfrak{G}}(\lambda) = \frac{1}{\prod_{i=1}^{m}(1-\lambda^{d_i})}, \text{ if } l = m,$$

or

(45b) $$\Phi_{\mathfrak{G}}(\lambda) = \frac{1+\sum_{j=l+1}^{m}\lambda^{d_j}}{\prod_{i=1}^{m}(1-\lambda^{d_i})}, \text{ if } l > m.$$

(This is easily verified by expanding (45a) and (45b) in powers of λ and comparing with (44b).)

Some examples will make this clear.

1. For the group $\mathfrak{G}_1$ of Part I, $f_1 = \theta$ and $f_2 = \varphi$ form a good polynomial basis, with degrees $d_1 = 8$, $d_2 = 24$. Indeed, from Th. 3a and Eq. (5),

$$\mathcal{I}(\mathfrak{G}_1) = \mathbb{C}[\theta,\varphi] \text{ and } \Phi_{\mathfrak{G}_1}(\lambda) = \frac{1}{(1-\lambda^8)(1-\lambda^{24})}.$$

2. For the group $\mathfrak{G}_4$ defined in Section B, $f_1 = x^2$, $f_2 = y^2$, $f_3 = xy$ is a good polynomial basis, with $d_1 = d_2 = d_3 = 2$. The invariants can be described as

(46) $$\mathcal{I}(\mathfrak{G}_3) = \mathbb{C}[x^2,y^2] \oplus xy\,\mathbb{C}[x^2,y^2].$$

In words, any invariant can be written uniquely as a polynomial in x^2 and y^2 plus xy times another such polynomial. E.g.,

$$(x+y)^4 = (x^2)^2 + 6x^2y^2 + (y^2)^2 + xy(4x^2+4y^2).$$

The Molien series is

$$\Phi_{\mathfrak{G}_3}(\lambda) = \frac{1}{2}\left\{\frac{1}{(1-\lambda)^2} + \frac{1}{(1+\lambda)^2}\right\} = \frac{1+\lambda^2}{(1-\lambda^2)^2}$$

in agreement with (45b) and (46). The single syzygy is $x^2 \cdot y^2 = (xy)^2$. Note that $f_1 = x^2$, $f_2 = xy$, $f_3 = y^2$ is not a good polynomial basis, for the invariant y^4 is not in the set $\mathbb{C}[x^2,xy] \oplus y^2\mathbb{C}[x^2,xy]$.

3. In [41] we gave an example, arising from coding theory, of a group $\mathfrak{G}_5$ (say) of 4×4 matrices with order 128, for which the Molien series is

$$\Phi_{\mathfrak{G}_5}(\lambda) = \frac{1+\lambda^8+\lambda^{10}+\lambda^{18}}{(1-\lambda^2)(1-\lambda^4)(1-\lambda^8)^2}.$$

A good polynomial basis for the invariants of this group consists of free invariants f_1, f_2, f_3, f_4, of degrees 2, 4, 8, 8, and transient invariants $f_4 = p$ (deg 8), $f_5 = q$ (deg 10), and $f_6 = pq$ (deg 18). Then if B denotes $\mathbb{C}[f_1,f_2,f_3,f_4]$, we have

$$\mathcal{I}(\mathfrak{G}_5) = B \oplus pB \oplus qB \oplus pqB.$$

There are syzygies expressing p^2 and q^2 in terms of $f_1,\ldots,f_6$.

4. A more complicated example, also arising from coding theory, is described in [**48**]. This is a group of 4×4 matrices, with order 336, having Molien series

$$\frac{1+\lambda^8+\lambda^{10}+\lambda^{12}+\lambda^{16}+\lambda^{18}+\lambda^{20}+\lambda^{28}}{(1-\lambda^4)(1-\lambda^6)(1-\lambda^8)(1-\lambda^{14})}.$$

A good polynomial basis is given in [**48**]. Fortunately the following result holds.

THEOREM 14 (Hochster and Eagon [**31**, Prop. 13]; independently proved by Dade [**19**]). *A good polynomial basis exists for the invariants of any finite group of complex $m \times m$ matrices.* (The proof is too complicated to give here.)

So we know that for any group the Molien series can be put into the standard form of Eqs. (45a), (45b) (with denominator consisting of a product of m factors $(1-\lambda^{d_i})$ and numerator consisting of a sum of powers of λ with positive coefficients); and that a good polynomial basis Eqs. (44a), (44b) can be found whose degrees match the powers of λ occurring in the Molien series.

On the other hand, the converse is not true. It is not always true that when the Molien series has been put into the form (45a), (45b) (by cancelling common factors and multiplying top and bottom by new factors), then a good polynomial basis for $\mathcal{I}(\mathfrak{G})$ can be found whose degrees match the powers of λ in $\Phi(\lambda)$. This is shown by the following example, due to Stanley [**71**].

Let $\mathfrak{G}_6$ be the group of order 8 generated by the matrices

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & i \end{pmatrix}.$$

The Molien series is

$$\Phi_{\mathfrak{G}_6}(\lambda) = \frac{1}{(1-\lambda^2)^3} \tag{47}$$

$$= \frac{1+\lambda^2}{(1-\lambda^2)^2(1-\lambda^4)}. \tag{48}$$

A good polynomial basis exists corresponding to Eq. (48), namely

$$\mathcal{I}(\mathfrak{G}_6) = \mathbb{C}[x^2, y^2, z^4] \oplus xy\, \mathbb{C}[x^2, y^2, z^4],$$

but there is no good polynomial basis corresponding to (47).

The problem of finding which forms of $\Phi(\lambda)$ correspond to a good polynomial basis and which do not remains unsolved in general.

However, one important special case has been solved. Shephard and Todd [**68**] have characterized those groups for which (44a) and (45a) hold, i.e., for which a good polynomial basis exists consisting only of algebraically independent invariants. These are the groups known as unitary groups generated by reflections. A complete list of the 37 irreducible groups of this type is given in [**68**].

IV. APPLICATIONS

We begin with a further example of the use of invariant theory to obtain results about weight enumerators. This illustrates the general plan of attack described in Section H of Part I in a situation where it is rather difficult to find a good polynomial basis. Other examples may be found in [**41**], [**49**] and [**50**]. It is worth mentioning that some of these examples use infinite groups and relative rather than absolute invariants.

A. Complete weight enumerator of a Ternary Self-Dual Code. Let $\mathscr{C}$ be an $[n, \frac{1}{2}n, d]$ self-dual code over F_3 which contains some codeword with no zeros. By suitably multiplying columns by -1

(which doesn't change the error-correcting ability of the code) we can assume that $\mathscr{C}$ contains the codeword $\mathbf{1} = 111\ldots1$.

Let A_{ijk} be the number of codewords in $\mathscr{C}$ containing i 0's, j 1's and k 2's (where $i+j+k=n$). Then the *complete weight enumerator* of $\mathscr{C}$ is defined to be

$$V(x,y,z) = \sum_{i,j,k} A_{ijk} x^i y^j z^k = \sum_{\mathbf{u}\in\mathscr{C}} x^{s_0(\mathbf{u})} y^{s_1(\mathbf{u})} z^{s_2(\mathbf{u})},$$

where $s_i(\mathbf{u})$ is the number of components of u that are equal to i. For example, the complete weight enumerator of the Golay code C_{11} (normalized to contain $\mathbf{1}$) is

$$\begin{aligned} V(x,y,z) = {} & x^{12}+y^{12}+z^{12}+22(x^6y^6+x^6z^6+y^6z^6) \\ & +220(x^6y^3z^3+x^3y^6z^3+x^3y^3z^6). \end{aligned} \tag{49}$$

The complete weight enumerator gives more information about a code than the weight enumerator $W(x,y)$ does. Of course the latter can be obtained from the equation $W(x,y) = V(x,y,y)$.

The goal of this section is to characterize the complete weight enumerator of $\mathscr{C}$ by proving:

THEOREM 15 ([**70**]). *If $V(x,y,z)$ is the complete weight enumerator of a self-dual code over F_3 which contains* $\mathbf{1}$, *then*

$$V(x,y,z) \in \mathbb{C}[\alpha_{12}, \beta_6^2, \delta_{36}] \oplus \beta_6\gamma_{18}\mathbb{C}[\alpha_{12}, \beta_6^2, \delta_{36}]$$

(i.e., *$V(x,y,z)$ can be written uniquely as a polynomial in α_{12}, β_6^2, δ_{36} plus $\beta_6\gamma_{18}$ times another such polynomial*), *where*

$$\begin{aligned} \alpha_{12} &= a(a^3+8p^3), \\ \beta_6 &= a^2-12b, \\ \gamma_{18} &= a^6-20a^3p^3-8p^6, \\ \delta_{36} &= p^3(a^3-p^3)^3, \end{aligned}$$

and

$$\begin{aligned} a &= x^3+y^3+z^3, \\ p &= 3xyz, \\ b &= x^3y^3+x^3z^3+y^3z^3. \end{aligned}$$

Note that $\gamma_{18}^2 = \alpha_{12}^3 - 64\delta_{36}$. (The subscript of a polynomial gives its degree.)

Proof. The proof follows the two stages described in Section H of Part I.

Stage I Let a typical codeword $\mathbf{u}\in\mathscr{C}$ contain a 0's, b 1's, and c 2's. Then since $\mathscr{C}$ is self-dual and contains $\mathbf{1}$

$$\begin{aligned} &\mathbf{u}\cdot\mathbf{u} = 0 \pmod 3 \Rightarrow 3\,|\,(b+c), \\ &\mathbf{u}\cdot\mathbf{1} = 0 \pmod 3 \Rightarrow 3\,|\,(b-c) \Rightarrow 3\,|\,b \text{ and } 3\,|\,c, \\ &\mathbf{1}\cdot\mathbf{1} = 0 \pmod 3 \Rightarrow 3\,|\,(a+b+c) \Rightarrow 3\,|\,a \end{aligned}$$

(where $a\,|\,b$ means "a divides b"). Therefore $V(x,y,z)$ is invariant under the transformations

$$\begin{pmatrix} \omega & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \omega \end{pmatrix}, \quad \omega = e^{2\pi i/3}.$$

Also $-\mathbf{u}$ contains a 0's, c 1's, b 2's, and $1+\mathbf{u}$ contains c 0's, a 1's, b 2's. Therefore $V(x,y,z)$ is

invariant under

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix},$$

i.e., under any permutation of its arguments.

Finally, from the MacWilliams' theorem for complete weight enumerators ([**42**] or [**43**]), $V(x, y, z)$ is invariant under

$$M_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{pmatrix}.$$

These 6 matrices generate a group $\mathfrak{G}_7$, of order 2592, consisting of 1944 matrices of the type

$$s^\nu \begin{pmatrix} 1 & & \\ & \omega^a & \\ & & \omega^b \end{pmatrix} M_3^e \begin{pmatrix} 1 & & \\ & \omega^c & \\ & & \omega^d \end{pmatrix}, \qquad s = e^{2\pi i/12},$$

and 648 matrices of the type

$$s^\nu \begin{pmatrix} 1 & & \\ & \omega^a & \\ & & \omega^b \end{pmatrix} P,$$

where $0 \leqq \nu \leqq 11$, $0 \leqq a, b, c, d \leqq 2$, $e = 1$ or 3, and P is any 3×3 permutation matrix.

Thus Stage I is completed: the assumptions about the code imply that $V(x, y, z)$ is invariant under the group $\mathfrak{G}_7$.

Stage II consists of showing that the ring of invariants of $\mathfrak{G}_7$ is equal to $\mathbf{C}[\alpha_{12}, \beta_{12}, \delta_{36}] \oplus \gamma_{24}\mathbf{C}[\alpha_{12}, \beta_{12}, \delta_{36}]$. First, since we have a list of the matrices in $\mathfrak{G}_7$, it is a straightforward hand calculation to obtain the Molien series, Eq. (3). As usual everything collapses and the final expression is

$$\Phi_{\mathfrak{G}_7}(\lambda) = \frac{1 + \lambda^{24}}{(1 - \lambda^{12})^2 (1 - \lambda^{36})}.$$

This suggests the degrees of a good polynomial basis that we should look for.

Next, $\mathfrak{G}_7$ is generated by J_3, M_3, and all permutation matrices P. Obviously the invariants must be symmetric functions of x, y, z having degree a multiple of 3. So we take the algebraically independent symmetric functions a, p, b, and find functions of them which are invariant under J_3 and M_3. For example, β_6 is invariant under J_3, but is sent into $-\beta_6$ by M_3. We denote this by writing

$$\beta_6 \overset{J_3}{\leftrightarrow} \beta_6, \qquad \beta_6 \overset{M_3}{\leftrightarrow} -\beta_6.$$

Therefore β_6^2 is an invariant. Again

$$a \overset{M_3}{\leftrightarrow} \frac{1}{\sqrt{3}}(a + 2p) \overset{J_3}{\rightarrow} \frac{1}{\sqrt{3}}(a + 2\omega p) \overset{M_3}{\leftrightarrow} \frac{i}{\sqrt{3}}(a + 2\omega^2 p),$$

so another invariant is $\alpha_{12} = a(a + 2p)(a + 2\omega p)(a + 2\omega^2 p) = a(a^3 + 8p^3)$. Again

$$\gamma_{18} \overset{J_3}{\leftrightarrow} \gamma_{18}, \qquad \gamma_{18} \overset{M_3}{\leftrightarrow} -\gamma_{18},$$

so $\beta_6\gamma_{18}$ is an invariant. Finally

$$p \overset{M_3}{\leftrightarrow} \frac{1}{\sqrt{3}}(a-p) \overset{J_3}{\rightarrow} \frac{1}{\sqrt{3}}(a-\omega p) \overset{M_3}{\leftrightarrow} \frac{s}{\sqrt{3}}(a-\omega^2 p)$$

gives the invariant

$$\delta_{36} = p^3(a-p)^3(a-\omega p)^3(a-\omega^2 p)^3 = p^3(a^3-p^3)^3.$$

The syzygy $\gamma_{18}^2 = \alpha_{12}^3 - 64\delta_{36}$ is easily verified, and one can show that α_{12}, β_6^2, δ_{36} are algebraically independent. Thus $f_1 = \alpha_{12}$, $f_2 = \beta_6^2$, $f_3 = \delta_{36}$, $f_4 = \beta_6\gamma_{18}$ is a good polynomial basis for $\mathscr{I}(\mathfrak{G}_7)$, and the theorem is proved. Q.E.D.

REMARK. Without the assumption that the code contains the all-ones vector, the theorem (due to R. J. McEliece) becomes much more complicated (see [**41**, §4.7], [**50**]).

Applications of Theorem 15. For the ternary Golay code (Eq. (49)) $V = \frac{1}{6}(5\alpha_{12} + \beta_6^2)$. For Pless's [24, 12, 9] symmetry code ([**56**], [**57**]),

$$V = \frac{67}{144}\alpha_{12}^2 + \frac{1}{8}\alpha_{12}\beta_6^2 + \frac{1}{432}\beta_6^4 + \frac{11}{27}\beta_6\gamma_{18}.$$

The complete weight enumerators of the symmetry codes of lengths 36, 48 and 60 have also been obtained with the help of Th. 15 (see [**50**]).

B. The nonexistence of certain very good codes. One application of Th. 3c was given in Section F of Part I, where it was used to determine the weight enumerator of a certain code. Other applications of this type may be found for example in [**47**].

A different type of application of Th. 3c is to show that certain codes with high minimum distance do not exist. The idea is to assume that the code does exist, and then to use Th. 3c to show that one of the coefficients in the weight enumerator is negative. But this is obviously impossible; therefore the code does not exist.

To explain the family of codes to which we shall apply this method, consider first the [7, 4, 3] Hamming code C_5 and the [23, 12, 7] Golay code C_7. These two codes have several properties in common, besides being quadratic residue codes as was mentioned earlier. For example they are both perfect codes. An $[n, k, d]$ code over F_q is called *perfect* if every vector is within Hamming distance $[\frac{1}{2}(d-1)]$ of some codeword. Perfect codes are very rare, and in fact Tietäväinen and van Lint (see [**72**], [**38**]) have given a complete list of all perfect codes. In particular, there is no binary perfect code with $d > 7$, and so the sequence $C_5, C_7, \ldots$ of binary perfect codes stops at C_7.

But there is another way of continuing this sequence. The extended codes, that is, the [8, 4, 4] code C_6 and the [24, 12, 8] code C_8, are both binary self-dual codes with all weights divisible by 4 and having minimum distance $d = 4[n/24] + 4$. This is the highest possible minimum distance for such a code, as the following argument shows.

Let $\mathscr{C}$ be any $[n, \frac{1}{2}n, d]$ binary self-dual code with all weights divisible by 4, and having weight enumerator $W(x, y)$. By Th. 3c, $W(x, y)$ is a polynomial in θ and φ and therefore can be written as

$$W(x,y) = \sum_{i=0}^{\mu} a_i \theta^{j-3i}\varphi^i, \tag{50}$$

where $n = 8j = 24\mu + 8\nu$, $\nu = 0, 1$ or 2.

Suppose the $\mu + 1 = [n/24] + 1$ coefficients a_i in (50) are chosen so that

$$\begin{aligned} W(x,y) &= x^n + A_{4\mu+4}x^{n-4\mu-4}y^{4\mu+4} + \cdots \\ &= W(x,y)^* \qquad \text{(say).} \end{aligned} \tag{51}$$

I.e., the a_i are chosen so that $W(x, y)$ has as many leading coefficients as possible equal to zero. An example of this was given in Section F of Part I. This determines the a_i and A_j uniquely. The resulting

W^* is the weight enumerator of that self-dual code with the greatest minimum weight we could hope to attain, and is called an *extremal* weight enumerator.

If a code exists with weight enumerator W^*, it has minimum distance $d^* = 4\mu + 4$, unless $A_{4\mu+4}$ is accidentally zero, in which case $d^* \geqq 4\mu + 8$.

But it can be shown [**47**] that $A_{4\mu+4}$, the number of codewords of minimum nonzero weight, is equal to:

$$\binom{n}{5}\binom{5\mu-2}{\mu-1}\Big/\binom{4\mu+4}{5}, \qquad \text{if } n = 24\mu,$$

$$\frac{1}{4}n(n-1)(n-2)(n-4)\frac{(5\mu)!}{\mu!(4\mu+4)!}, \qquad \text{if } n = 24\mu + 8,$$

$$\frac{3}{2}n(n-2)\frac{(5\mu+2)!}{\mu!(4\mu+4)!}, \qquad \text{if } n = 24\mu + 16,$$

and is never zero. This proves

THEOREM 16 ([**47**]). *The minimum distance of a binary self-dual code of length n with all weights divisible by* 4 *is at most* $4[n/24]+4$.

So it is natural to study the sequence of self-dual codes with weights divisible by 4 and minimum distance actually equal to $4[n/24]+4$. Such codes are known to exist for $n \leqq 48$ and for a few larger values of n (see [**42**]). Length $n = 72$ is the most important open case (see [**69**]). But the next theorem shows that this sequence of codes (like that of perfect codes) is finite.

In fact it turns out that the second coefficient in (51), $A_{4\mu+8}$, is negative if n is large (above about 3712), and so a self-dual code with weight enumerator W^* does not exist for large n. Furthermore, one can show that no self-dual code can even have minimum distance within a constant of $n/6$, if n is sufficiently large:

THEOREM 17 ([**46**]). *Let b be any constant. Suppose the a_i in* (50) *are chosen so that*

$$W(x, y) = x^n + A_{4d}x^{n-4d}y^{4d} + \cdots,$$

where $d \geqq n/6 - b$. Then one of the coefficients A_i is negative, for all sufficiently large n. So a binary self-dual code of length n, weights divisible by 4, *and minimum weight d does not exist for all sufficiently large n.*

On the other hand it is known that binary self-dual codes with weights divisible by 4 do meet the Gilbert–Varshamov bound [**44**].

Similar results hold for ternary self-dual codes and for certain types of lattice sphere packings (see [**46**]).

V. CONCLUSIONS

We have attempted to show how invariant theory has been used to solve problems in coding theory. There are two stages in a typical application of this technique, which is potentially of much wider application. Stage I: convert assumptions about the problem (e.g., the code) into algebraic constraints on polynomials (e.g., on the weight enumerator of the code). Stage II: use invariant theory to find all possible polynomials satisfying these constraints.

Some unsolved problems are (i) what is the greatest n for which the upper bound of Th. 16 is attained? In particular, does such a code exist with $n = 72$ (see [**69**])? (ii) An unsolved question from [**41**]: characterize the biweight enumerator of a binary self-dual code with all weights divisible by 4. (iii) For a given group of matrices, which forms of the Molien series $\Phi(\lambda)$ correspond to a good polynomial basis? (iv) Given two different groups of matrices, $\mathfrak{G}$ and $\mathfrak{H}$, which are both representations of the same abstract group, what is the relationship between $\mathcal{I}(\mathfrak{G})$ and $\mathcal{I}(\mathfrak{H})$?

Acknowledgments: This article is based on several joint papers **[41]**, **[46]–[49]** with F. J. MacWilliams, C. L. Mallows and A. M. Odlyzko. I am very grateful for their collaboration. I should also like to thank W. L. Edge, J. McKay, M. K. Sain and R. Stanley for their helpful comments. Some of the calculations were verified with the help of the ALTRAN system for rational function manipulation [**13**], [**30**], and the GRAPPA group theory analysis program [**21**]. There is some overlap with the earlier survey [**70**], although the present paper is written from a different point of view and for a different audience.

References

1. V. K. Agrawala and J. G. Belinfante, An Algorithm for Computing SU(n) Invariants, BIT II (1971) 1–15.

2. A. C. Aitken, Determinants and Matrices, Oliver and Boyd, 9-th ed., Edinburgh, 1959.

3. E. F. Assmus, Jr., and H. F. Mattson, Jr., New 5-designs, J. Combinatorial Theory, 6 (1969) 122–151.

4. ———, and ———, Coding and combinatorics, SIAM Review, 16 (1974) 349–388.

5. E. F. Assmus, Jr., H. F. Mattson, and R. Turyn, Cyclic codes, Air Force Cambridge Res. Lab., Report AFCRL-65-322, Bedford, Mass., 1965.

6. ———, ———, and ———, Cyclic codes, Air Force Cambridge Res. Lab., Report AF19(604)–8516, Bedford, Mass., 1966.

7. E. R. Berlekamp, Algebraic Coding Theory, McGraw–Hill, New York, N.Y., 1968.

8. ———, Key papers in the development of coding theory, IEEE Press, New York, 1974.

9. E. R. Berlekamp, F. J. MacWilliams, and N. J. A. Sloane, Gleason's theorem on self-dual codes, IEEE Trans. Info. Theory, IT-18 (1972) 409–414.

10. I. F. Blake, Algebraic Coding Theory: History and Development, Dowden, Stroudsburg, Pa., 1973.

11. N. Bourbaki, Groupes et Algèbres de Lie, Ch. 4, 5 & 6, Hermann, Paris, 1968.

12. M. Broué and M. Enguehard, Polynômes des poids de certains codes et fonctions thêta de certains réseaux, Ann. Scient. Ec. Norm. Sup., 5 (1972) 157–181.

13. W. S. Brown, ALTRAN User's Manual, 3rd ed., Bell Laboratories, N. J. 1974.

14. W. Burnside, Theory of Groups of Finite Order, 2nd ed. Dover, New York, 1955.

15. J. H. Conway, A group of order 8, 315, 553, 613, 086, 720, 000, Bull. London Math. Soc., 1 (1969) 79–88.

16. ———, A characterization of Leech's lattice, Invent. Math., 7 (1969) 137–142.

17. ———, Groups, lattices, and quadratic forms, in: Computers in Algebra and Number Theory, SIAM–AMS Proc. IV, Amer. Math. Soc., 1971, 135–139.

18. E. C. Dade, Answer to a question of R. Brauer, J. Algebra, 1 (1964) 1–4.

19. E. C. Dade, letter to the author, 23 April 1972.

20. J. Dieudonné and J. B. Carrell, Invariant Theory, Old and New, Academic Press, New York, 1971.

21. L. A. Dimino, GRAPPA Basic User's Manual, Bell Laboratories, Murray Hill, N.J. 1972.

22. P. Doubilet, G.-C. Rota, and J. Stein, On the foundations of combinatorial theory: IX Combinatorial Methods in Invariant Theory, Studies in Appl. Math., 53 (1974) 185–216.

23. H. Dym and H. P. McKean, Fourier Series and Integrals, Academic Press, New York, 1972.

24. W. Feit, A self-dual even (96, 48, 16) code, IEEE Trans. Info. Theory IT-20 (1974) 136–138.

25. C. S. Fisher, The death of a mathematical theory: A study in the sociology of knowledge, Arch. Hist. Exact Sci., 3 (1967) 136–159.

26. A. M. Gleason, Weight polynomials of self-dual codes and the MacWilliams identities, in: Actes Congrès Internl. de Mathématique, 3 (1970) (Gauthier-Villars, Paris 1971) 211–215.

27. J.-M. Goethals, On the Golay perfect binary code, J. Combin. Theory, 11(A) (1971) 178–186.

28. M. J. E. Golay, Notes on digital coding, Proc. IEEE, 37 (1949) 657.

29. ———, Binary Coding, IEEE Trans. Info. Theory, 4 (1954) 23–28.

30. A. D. Hall, Jr., The ALTRAN system for rational function manipulation — A survey, Comm. Assoc. Comput. Mach., 14 (1971) 517–521.

31. M. Hochster and J. A. Eagon, Cohen–Macaulay rings, invariant theory, and the generic perfection of determinantal loci, Amer. J. Math., 93 (1971) 1020–1058.

32. N. Jacobson, Lectures in Abstract Algebra, Vol. 3, Van Nostrand, Princeton, N.J., 1964.

33. M. G. Kendall and A. Stuart, The Advanced Theory of Statistics, Vol. 1, Hafner, New York, 1969.

34. J. Leech and N. J. A. Sloane, Sphere packings and error-correcting codes, Canad. J. Math., 23 (1971) 718–745.

35. V. K. Leontjev, Spectra of Linear Codes, Third International Symposium on Information Theory, Tallinn, Estonia, June 1973, Abstracts of Papers, Part II, pp. 102–106.

36. N. Levinson, Coding Theory: A counterexample to G. H. Hardy's conception of applied mathematics, this MONTHLY, 77 (1970) 249–258.

37. J. H. van Lint, Coding theory, Lecture Notes in Math., 201 (Springer, Berlin, 1971).

38. ———, Recent results on perfect codes and related topics, pp. 163–183 of *Combinatorics*, edited by M. Hall, Jr., and J. H. van Lint, Reidel Publishing Co., Dordrecht, Holland, 1975.

39. D. E. Littlewood, A University Algebra, 2nd ed., Dover, New York, 1970.

40. F. J. MacWilliams, A theorem on the distribution of weights in a systematic code, Bell Syst. Tech. J., 42 (1963) 79–84.

41. F. J. MacWilliams, C. L. Mallows, and N. J. A. Sloane, Generalization of Gleason's theorem on weight enumerators of self-dual codes, IEEE Trans. Info. Theory, IT-18 (1972) 794–805.

42. F. J. MacWilliams and N. J. A. Sloane, The Theory of Error-correcting Codes, North Holland, Amsterdam, 1977, to appear.

43. F. J. MacWilliams, N. J. A. Sloane, and J. M. Goethals, The MacWilliams identities for nonlinear codes, Bell System Tech. J., 51 (1972) 803–819.

44. F. J. MacWilliams, N. J. A. Sloane, and J. G. Thompson, Good self-dual codes exist, Discrete Math., 3 (1972) 153–162.

45. ———, ———, and ———, On the existence of a projective plane of order 10, J. Combinatorial Theory, 14A (1973) 66–78.

46. C. L. Mallows, A. M. Odlyzko, and N. J. A. Sloane, Upper bounds for modular forms, lattices, and codes, J. Algebra, 36 (1975) 68–76.

47. C. L. Mallows and N. J. A. Sloane, An upper bound for self-dual codes, Information and Control, 22 (1973) 188–200.

48. ——— and ———, On the invariants of a linear group of order 336, Proc. Camb. Phil. Soc., 74 (1973) 435–440.

49. ——— and ———, Weight enumerators of self-orthogonal codes, Discrete Math., 9 (1974) 391–400.

50. C. L. Mallows, V. Pless, and N. J. A. Sloane, Self-dual codes over GF(3), SIAM J. Applied Math., 31 (December 1976).

51. G. A. Miller, H. F. Blichfeldt and L. E. Dickson, Theory and Applications of Finite Groups, Dover, New York, 1961.

52. T. Molien, Über die Invarianten der linearen Substitutionsgruppe, Sitzungsber. König. Preuss. Akad. Wiss., (1897) 1152–1156.

53. D. Mumford, Geometric Invariant Theory, Springer, New York, 1965.

54. E. Noether, Der Endlichkeitssatz der Invarianten endlicher Gruppen, Math. Ann., 77 (1916) 89–92.

55. W. W. Peterson and E. J. Weldon, Jr., Error-Correcting Codes, 2nd ed., MIT Press, Cambridge, Mass., 1972.

56. V. Pless, On a new family of symmetry codes and related new five-designs, Bull. Amer. Math. Soc., 75 (1969) 1339–1342.

57. ———, Symmetry Codes over GF(3) and New Five-Designs, J. Combinatorial Theory, 12 (1972) 119–142.

58. ———, A classification of self-orthogonal codes over GF(2), Discrete Math., 3 (1972) 209–246.

59. ———, Self-dual codes, Proc. 5th S. -E. Conference on Combinatorics, Graph Theory, and Computing, Utilitas Math., (1974) 111–124.

60. V. Pless and J. N. Pierce, Self-dual codes over GF(q) satisfy a modified Varshamov bound, Information and Control, 23 (1973) 35–40.

61. V. Pless and N. J. A. Sloane, Binary self-dual codes of length 24, Bull. Amer. Math. Soc., 80 (1974) 1173–1178.

62. ——— and ———, On the classification and enumeration of self-dual codes, J. Combinatorial Theory, 18A (1975) 313–335.

63. S. J. Rallis, New and old results in invariant theory with applications to arithmetic groups, pp. 443–458 of *Symmetric Spaces*, edited by W. M. Boothby and G. L. Weiss, Dekker, New York, 1972.

64. C. Reid, Hilbert, Springer–Verlag, New York, 1970.

65. G. -C. Rota, Combinatorial theory and invariant theory, lecture notes, Bowdoin College, Maine, Summer 1971.

66. J. -P. Serre, Représentations Linéaires des Groupes Finis, 2nd ed., Hermann, Paris, 1971.

67. E. P. Shaughnessy, Codes with simple automorphism groups, Arch. Math., 22 (1971) 459–466.

68. G. C. Shephard and J. A. Todd, Finite unitary reflection groups, Canad. J. Math., 6 (1954) 274–304.

69. N. J. A. Sloane, Is there a (72, 36) $d = 16$ self-dual code? IEEE Trans. Info. Theory, IT-19 (1973) 251.

70. ———, Weight enumerators of codes, pp. 115–142 of *Combinatorics*, M. Hall, Jr., and J. H. van Lint, editors, Reidel Publishing Co., Dordrecht, Holland, 1975.

71. R. P. Stanley, private communication.

72. A. Tietäväinen, On the existence of perfect codes over finite fields, SIAM J. Appl. Math., 24 (1973) 88–96.

73. H. Weyl, Invariants, Duke Math. J., 5 (1939) 489–502.

74. ———, The Classical Groups, Princeton University Press, Princeton, N.J., 1946.

ROOM 2C–363, BELL LABORATORIES, MURRAY HILL, N.J. 07974.

SINGULAR CARDINALS AND THE GENERALIZED CONTINUUM HYPOTHESIS

J. E. BAUMGARTNER AND K. PRIKRY

1. Introduction. By now, set theory has developed to a point where significant new results of an elementary nature are quite rare. Progress is usually made by means of sophisticated metamathematical techniques such as the method of ultrapowers and Cohen's theory of forcing and generic sets. It was by means of such methods that Silver [**11**] recently proved a remarkable theorem about the generalized continuum hypothesis. Rather surprisingly, it has turned out that Silver's theorem has a purely elementary proof (see [**2**]). The purpose of this paper is to give a proof of a typical case of Silver's theorem which is completely self-contained except for the few well-known facts outlined in the remainder of this section.

The rest of the paper is organized as follows:

Section 2 contains a brief history of the continuum problem. Section 3 is devoted to the proof of another important elementary theorem, due to Fodor, which will be necessary for the main proof. In section 4 we prove Silver's theorem that if $2^{\aleph_\alpha} = \aleph_{\alpha+1}$ for all $\alpha < \omega_1$, then $2^{\aleph_{\omega_1}} = \aleph_{\omega_1+1}$. Some generalizations and extensions of Silver's theorem are discussed in section 5.

We presume the reader knows that ordinal numbers may be defined in such a way that each ordinal coincides with the set of all smaller ordinal numbers. Thus for ordinal numbers α and β, the expressions "$\alpha < \beta$" and "$\alpha \in \beta$" are synonymous. The axiom of choice, which is assumed throughout this paper, implies that all the infinite cardinal numbers may be arranged in an increasing ordinal-indexed sequence $\aleph_0, \aleph_1, \aleph_2, \ldots, \aleph_\alpha, \ldots$. The first ordinal number of cardinality $\aleph_\alpha$ will be denoted by ω_α. Therefore, for example, ω_0 is the set of all nonnegative integers (finite ordinals), and ω_1 is the set of all ordinals which are countable or finite. We shall always denote ω_0 by ω.

If X is a set then the *power set* of X, the set of all subsets of X, will be denoted by $P(X)$. Of course, if X has cardinality $\aleph_\alpha$, then $P(X)$ has cardinality $2^{\aleph_\alpha}$. The set of all functions from a set X into a set Y will be denoted by Y^X. If X has cardinality $\aleph_\alpha$ and Y has cardinality $\aleph_\beta$, then Y^X has cardinality $\aleph_\beta^{\aleph_\alpha}$.

We will need the following easy facts about cardinal arithmetic:

(1) For all α, $2^{\aleph_\alpha} \geqq \aleph_{\alpha+1}$.

(2) For all α and β, $\aleph_\alpha \cdot \aleph_\beta = \max(\aleph_\alpha, \aleph_\beta)$.

(3) For all α and β, if $\alpha < \beta$ then $\aleph_\alpha^{\aleph_\beta} = \aleph_\beta^{\aleph_\beta}$ and $\aleph_\beta^{\aleph_\alpha} \leqq \aleph_\beta^{\aleph_\beta}$.

(4) For all α, $\aleph_\alpha^{\aleph_\alpha} = 2^{\aleph_\alpha}$.

A cardinal $\aleph_\alpha$ is said to be *regular*, provided that for any indexed family $\{S_i : i \in I\}$ of sets, if I and all the sets S_i have cardinality less than $\aleph_\alpha$, then so does $\cup\{S_i : i \in I\}$. It is easily seen that every cardinal of the form $\aleph_{\alpha+1}$ is regular, for if I and S_i $(i \in I)$ are as above, all with cardinality $\leqq \aleph_\alpha$, then $\cup\{S_i : i \in I\}$ has cardinality at most $\aleph_\alpha \cdot \aleph_\alpha = \aleph_\alpha$. One consequence of the regularity of $\aleph_1$ is the fact that if $\{\alpha_n : n \in \omega\}$ is a collection of elements of ω_1, then $\sup\{\alpha_n : n \in \omega\} < \omega_1$.

If $\aleph_\alpha$ is not regular, it is called *singular*. For example, $\aleph_\omega$ and $\aleph_{\omega_1}$ are singular cardinals.

An excellent reference for the material above is P. R. Halmos's book, *Naive Set Theory*.

2. Historical survey. Georg Cantor, the founder of set theory, was the first to consider the problem of determining the ordinal α for which $2^{\aleph_0} = \aleph_\alpha$ is true. Since $2^{\aleph_0}$ is the cardinality of the continuum of real numbers, this is called the *continuum problem*. Cantor was only able to show that $2^{\aleph_0} > \aleph_0$, so he made the natural conjecture that $2^{\aleph_0} = \aleph_1$; this is the *continuum hypothesis*. Cantor made several attempts to prove the continuum hypothesis, as did others, including David Hilbert, but none of the proofs has turned out to be satisfactory.

Further information about the continuum problem was provided in 1905 by Julius König [**9**] in the following theorem:

KÖNIG'S THEOREM. *Suppose* $\{\kappa_i\}_{i\in I}$ *and* $\{\lambda_i\}_{i\in I}$ *are indexed sets of (possibly finite) cardinal numbers such that* $\kappa_i < \lambda_i$ *for all* $i \in I$. *Then*

$$\sum_{i\in I} \kappa_i < \prod_{i\in I} \lambda_i.$$

(This theorem may seem to be very weak; try to strengthen it!)

Let us give two illustrative applications of König's theorem. First, let $I = \omega$ and let $\kappa_i = 1$ and $\lambda_i = 2$ for all i. Then

$$\sum_{i\in\omega} \kappa_i = \aleph_0, \qquad \text{and} \qquad \prod_{i\in\omega} \lambda_i = 2^{\aleph_0},$$

so $\aleph_0 < 2^{\aleph_0}$, which is Cantor's theorem.

Second, suppose $2^{\aleph_0} \geqq \aleph_\omega$. Let $I = \omega$ again, and for each $i \in \omega$ let $\kappa_i = \aleph_i$ and $\lambda_i = \aleph_\omega$. Then

$$\sum_{i\in\omega} \kappa_i = \aleph_\omega, \qquad \text{and} \qquad \prod_{i\in\omega} \lambda_i = \aleph_\omega^{\aleph_0} \leqq (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0}.$$

Hence $\aleph_\omega < 2^{\aleph_0}$. It follows that $2^{\aleph_0} \neq \aleph_\omega$. Moreover, the same argument shows that if α is the limit of a countable sequence of smaller ordinals, then $2^{\aleph_0} \neq \aleph_\alpha$.

The widespread belief that the continuum hypothesis should be provable or refutable in existing systems of set theory was given a shock in 1930 by Gödel's famous Incompleteness Theorem. Gödel's proof showed that for any (recursively) axiomatizable formal system of set theory[1], if the system itself is consistent, then there are sentences which are neither provable nor refutable (i.e., are *undecidable*) in the system. Of course, if the system is inconsistent then *every* sentence is provable. The examples of undecidable sentences yielded by Gödel's proof were mathematically rather unnatural, but suspicion began to grow that the continuum hypothesis might be undecidable in the usual systems of set theory.

Half of this suspicion was confirmed by Gödel [8] in 1939, when he showed that for many systems of set theory the continuum hypothesis could not be disproved unless the system itself was inconsistent. Let us take a precise example. The axiomatic system used most frequently is that of Zermelo and Fraenkel, which we denote by ZF. When the axiom of choice is added we denote the result by ZFC. Let GCH stand for the *generalized continuum hypothesis*, that is to say the assertion that $2^{\aleph_\alpha} = \aleph_{\alpha+1}$ for every ordinal α. Then Gödel's argument showed that if ZF is consistent then so is ZFC + GCH.

The remaining piece of the puzzle was found by Paul Cohen [4] in 1963. By means of his method of forcing and generic sets, he was able to show that if ZF is consistent, then so is ZF together with the negation of the axiom of choice, and so is ZFC + $2^{\aleph_0} \geqq \aleph_2$. Therefore, provided that ZF is consistent, the continuum hypothesis is undecidable in ZFC.

Shortly after Cohen's breakthrough, Solovay [12] showed that in Cohen's result $2^{\aleph_0}$ could take any value not excluded by König's theorem. Thus, for example, if ZF is consistent, then so is ZFC + $2^{\aleph_0} = \aleph_{14}$, so is ZFC + $2^{\aleph_0} = \aleph_{\omega+1}$, and so is ZFC + $2^{\aleph_0} = \aleph_{\omega_1}$.

Easton [5] extended these methods to treat the continuum problem for regular cardinals greater than $\aleph_0$ and to treat the continuum problem for many regular cardinals at once. Examples of his results are as follows: If ZF is consistent, then so is

$$\mathrm{ZFC} + 2^{\aleph_0} = \aleph_1 + 2^{\aleph_1} = \aleph_3 + 2^{\aleph_2} = \aleph_{105},$$

and so is

$$\mathrm{ZFC} + \text{"for all regular cardinals } \aleph_\alpha,\ 2^{\aleph_\alpha} = \aleph_{\alpha+2}\text{"}.$$

[1] Here we mean a *sufficiently strong* system of set theory. The technical criterion is that all recursive sets should be representable in it.

The continuum problem for singular cardinals cannot be treated by Easton's methods. This problem has become known as the *singular cardinals problem*, and many parts of it are still unsolved today. To take the simplest example, suppose $2^{\aleph_n} = \aleph_{n+1}$ for all $n \in \omega$. Does it follow that $2^{\aleph_\omega} = \aleph_{\omega+1}$? The answer is still unknown.[2]

It was therefore all the more surprising when Silver **[11]** recently proved that if $2^{\aleph_\alpha} = \aleph_{\alpha+1}$ for all $\alpha < \omega_1$, then $2^{\aleph_{\omega_1}} = \aleph_{\omega_1+1}$! The rest of this paper is devoted to an elementary proof of Silver's result which was found by the authors. (A similar proof has been found by Jensen.) Silver's result is actually much more general than is possible to indicate here. The interested reader should consult **[2]**, **[7]**, or **[11]**.

Silver's theorem has been generalized by Galvin and Hajnal **[7]**.

Jensen has also obtained a remarkable result about the singular cardinals problem. He showed recently that if one wants to prove the consistency of

$$(*)\qquad \text{ZFC} + \text{“for all } n \in \omega, 2^{\aleph_n} < \aleph_\omega \text{ and } 2^{\aleph_\omega} \geqq \aleph_{\omega+2}\text{”},$$

then it is necessary to assume *more* than the consistency of ZF! On the other hand, by assuming the consistency of ZFC together with certain "large cardinal" axioms, Magidor **[10]** has proved the consistency of (*).

3. Stationary sets and Fodor's theorem. This section contains a theorem of Fodor **[6]** which is fundamental for a large part of modern set theory. In view of its elementary nature, it is surprising that the theorem is seldom included in set theory textbooks.

A subset C of ω_1 is called *closed* provided that for every nonempty set $A \subseteq C$, if $\sup A < \omega_1$, then $\sup A \in C$. It is easy to see that a set is closed in this sense if and only if it is closed in the order topology on ω_1. It follows that the intersection of any number of closed sets is closed. (This is also trivial to verify directly.)

Let us call $C \subseteq \omega_1$ *unbounded* if for every $\alpha < \omega_1$ there is $\beta > \alpha$ such that $\beta \in C$. We shall be interested in sets which are both closed and unbounded. It is not true that the intersection of any number of closed unbounded sets is closed and unbounded, but some useful partial results can be obtained.

THEOREM 1. *Suppose that for each natural number $n \in \omega$, C_n is a closed unbounded set. Let C be the intersection of the sets C_n. Then C is also closed and unbounded.*

Proof. We have already observed that C must be closed. Suppose $\alpha < \omega_1$. We must find $\beta > \alpha$ so that $\beta \in C$. The ordinal β will be obtained as the supremum of an increasing sequence $\langle \beta_m : m \in \omega \rangle$ of countable ordinals defined by induction as follows. Let $\beta_0 = \alpha$. Now suppose β_m has been obtained. Since each C_n is unbounded there is an ordinal $\gamma(m, n) \in C_n$ such that $\beta_m < \gamma(m, n)$. Let $\beta_{m+1} = \sup\{\gamma(m, n): n \in \omega\}$. Since $\aleph_1$ is regular, $\beta_{m+1} < \omega_1$. Let $\beta = \sup\{\beta_m : m \in \omega\}$. Then $\beta < \omega_1$ also. We must show that for each n, $\beta \in C_n$. But since $\beta_m < \gamma(m, n) \leqq \beta_{m+1}$ for all m and n, it is clear that for each n, $\sup\{\gamma(m, n): m \in \omega\} = \sup\{\beta_m : m \in \omega\} = \beta$. Since C_n is closed, $\sup\{\gamma(m, n): m \in \omega\} \in C_n$. Hence $\beta \in C$, so C is unbounded. $\square$

Note that Theorem 1 does not remain true for intersections of uncountably many closed unbounded sets. For example, let $C_\alpha = \{\beta < \omega_1: \alpha \leqq \beta\}$ for each $\alpha < \omega_1$. Then each C_α is closed and unbounded, but the intersection of the sets C_α is empty. Nevertheless a theorem very close to this *is* true; this is one version of Fodor's theorem.

THEOREM 2. *For each $\alpha < \omega_1$ let C_α be a closed unbounded set. Let $C = \{\beta < \omega_1$: for all $\alpha < \beta$, $\beta \in C_\alpha\}$. Then C is closed and unbounded.*

[2] **Note added in proof:** M. Magidor has recently shown that the answer is negative, assuming the consistency of the existence of certain very large cardinal numbers.

Proof. Suppose $A \subseteq C$, A is nonempty, and $\sup A = \beta < \omega_1$. Fix $\alpha < \beta$, and let $A_\alpha = \{\gamma \in A : \alpha < \gamma\}$. Then $\sup A_\alpha = \sup A = \beta$ and A_α is nonempty. Since $A_\alpha \subseteq C_\alpha$ (by definition of C) and C_α is closed, $\beta \in C_\alpha$. Since α was arbitrary, this means $\beta \in C$. Hence C is closed.

Now suppose $\alpha < \omega_1$. We must find $\beta > \alpha$ such that $\beta \in C$. As in the proof of Theorem 1 we obtain β as the supremum of an increasing sequence $\langle \beta_n : n \in \omega \rangle$ of countable ordinals. Let $\beta_0 = \alpha$. Suppose β_n has been obtained. Since $\{\gamma : \gamma < \beta_n\}$ is countable, we may apply Theorem 1 to see that $\bigcap\{C_\gamma : \gamma < \beta_n\}$ is closed and unbounded. Choose $\beta_{n+1} > \beta_n$ so that $\beta_{n+1} \in \bigcap\{C_\gamma : \gamma < \beta_n\}$. Let $\beta = \sup\{\beta_n : n \in \omega\}$, and suppose $\delta < \beta$. We must show $\beta \in C_\delta$. Now for some n it must be true that $\delta < \beta_n$. But then we have $\{\beta_m : m > n\} \subseteq C_\delta$ by construction. Since C_δ is closed, $\sup\{\beta_m : m > n\} = \beta \in C_\delta$. □

REMARK. The set C of Theorem 2 is sometimes called the *diagonal intersection* of the sets C_α.

The rest of this section is devoted to recasting Theorems 1 and 2 in a more usable form.

A set $S \subseteq \omega_1$ is called *stationary* provided that $S \cap C$ is nonempty for every closed unbounded set C.

THEOREM 3. (a) *Every closed unbounded set is stationary.*

(b) *If S is stationary and C is closed and unbounded then $S \cap C$ is stationary.*

Proof. By Theorem 1. □

We leave it to the reader to verify that every stationary set is unbounded.

It may be helpful intuitively to think of closed unbounded sets as very large subsets of ω_1, and to think of stationary sets as fairly large subsets of ω_1. The following version of Theorem 1 may then be interpreted as saying that if ω_1 is decomposed into countably many pieces, then at least one of the pieces must be fairly large.

THEOREM 4. *Suppose that $\omega_1 = \bigcup\{S_n : n \in \omega\}$. Then at least one of the sets S_n is stationary.*

Proof. Suppose not. Then for each n there is a closed unbounded set C_n disjoint from S_n. By Theorem 1, $\bigcap\{C_n : n \in \omega\}$ is nonempty. But $\bigcap\{C_n : n \in \omega\}$ must be disjoint from $\bigcup\{S_n : n \in \omega\} = \omega_1$, a contradiction. □

The following remarkable theorem is the most frequently stated version of Fodor's theorem.

THEOREM 5. *Suppose S is stationary and f is a function such that $f(\alpha) < \alpha$ for all $\alpha \in S$. Then there is a stationary subset S' of S on which f is constant.*

Proof. Suppose not. Then for each $\alpha < \omega_1$ there is a closed unbounded set C_α such that $f(\beta) \neq \alpha$ for all $\beta \in C_\alpha$. Let C be the diagonal intersection of the C_α as in Theorem 2. Then C is closed and unbounded so $S \cap C$ is nonempty. Let $\beta \in S \cap C$ and suppose $f(\beta) = \alpha$. Then by definition of C_α, $\beta \notin C_\alpha$. But since $\alpha < \beta$ and $\beta \in C$, $\beta \in C_\alpha$ and this is a contradiction. □

REMARK. The observation that if $f(\alpha) < \alpha$ for all α such that $1 \leqq \alpha < \omega_1$, then f is constant on an uncountable set, was made as early as 1929 by Alexandroff and Urysohn [**1**]. The definition of stationary sets is due to G. Bloch [**3**].

Finally we state the consequence of Fodor's theorem which will be needed in the proof of Silver's theorem.

THEOREM 6. *Suppose S is a stationary subset of ω_1 and f is a function such that for all $\alpha \in S$, $f(\alpha) < \omega_\alpha$. Then there is an ordinal $\gamma < \omega_1$ and a stationary set $S' \subseteq S$ such that for all $\alpha \in S'$, $f(\alpha) < \omega_\gamma$.*

Proof. Let C be the set of limit ordinals less than ω_1, i.e., the set of ordinals which are neither 0 nor of the form $\beta + 1$ for some β. It is easy to see that C is closed and unbounded. If $\alpha \in S \cap C$ then $\omega_\alpha = \sup\{\omega_\beta : \beta < \alpha\}$ so there is some $\beta < \alpha$ such that $f(\alpha) < \omega_\beta$. Let $g(\alpha)$ be the least such β. Since

$S \cap C$ is stationary by Theorem 3(b), we may apply Theorem 5 to obtain $S' \subseteq S \cap C$ and γ such that S' is stationary and $g(\alpha) = \gamma$ for every $\alpha \in S'$. But then $f(\alpha) < \omega_\gamma$ for every $\alpha \in S'$. $\square$

4. Silver's theorem.

THEOREM 7. (Silver) *Suppose that for every* $\alpha < \omega_1$, $2^{\aleph_\alpha} = \aleph_{\alpha+1}$. *Then* $2^{\aleph_{\omega_1}} = \aleph_{\omega_1+1}$.

Proof. We must show that $P(\omega_{\omega_1})$ has cardinality $\aleph_{\omega_1+1}$.

For each $\alpha < \omega_1$, let $\langle A^\alpha_\xi : \xi < \omega_{\alpha+1}\rangle$ be a listing without repetitions of all the elements of $P(\omega_\alpha)$. Since $2^{\aleph_\alpha} = \aleph_{\alpha+1}$, such a listing exists.

For each $A \subseteq \omega_{\omega_1}$, define f_A on ω_1 by letting $f_A(\alpha) = \xi$ if and only if $A \cap \omega_\alpha = A^\alpha_\xi$. Note that if $A \neq B$ then for some $\alpha < \omega_1$, $A \cap \omega_\alpha \neq B \cap \omega_\alpha$, and it follows that for all $\beta \geqq \alpha$, $f_A(\beta) \neq f_B(\beta)$. Hence $\{\alpha : f_A(\alpha) = f_B(\alpha)\}$ is bounded.

Now we define a relation R between members of $P(\omega_{\omega_1})$ by letting ARB if and only if $\{\alpha : f_A(\alpha) < f_B(\alpha)\}$ is stationary.

LEMMA 1. *If* A, $B \in P(\omega_{\omega_1})$ *and* $A \neq B$ *then either* ARB *or* BRA.

Proof. Clearly

$$\omega_1 = \{\alpha : f_A(\alpha) < f_B(\alpha)\} \cup \{\alpha : f_A(\alpha) = f_B(\alpha)\} \cup \{\alpha : f_A(\alpha) > f_B(\alpha)\},$$

so by Theorem 4 at least one of these three sets must be stationary. Moreover, it can't be the middle one since $\{\alpha : f_A(\alpha) = f_B(\alpha)\}$ is bounded. $\square$

For the rest of the proof, let us assume that $P(\omega_{\omega_1})$ has cardinality $> \aleph_{\omega_1+1}$. We shall obtain a contradiction.

LEMMA 2. *There exists* $B \subseteq \omega_{\omega_1}$ *such that* $\{A : ARB\}$ *has cardinality at least* $\aleph_{\omega_1+1}$.

Proof. Let X be a subset of $P(\omega_{\omega_1})$ of cardinality $\aleph_{\omega_1+1}$. If there is $B \in X$ which satisfies the lemma then we are done, so suppose not. For each $B \in X$, let $R^{-1}(B) = \{A : ARB\}$. Let $Y = \cup\{R^{-1}(B) : B \in X\}$. Then Y is a union of $\aleph_{\omega_1+1}$ sets, each of which has cardinality at most $\aleph_{\omega_1}$, so Y has cardinality at most $\aleph_{\omega_1+1} \cdot \aleph_{\omega_1} = \aleph_{\omega_1+1}$. Since $P(\omega_{\omega_1})$ has cardinality $> \aleph_{\omega_1+1}$, there is $B \in P(\omega_{\omega_1})$ such that $B \notin Y$. Now if $A \in X$ then $B \notin R^{-1}(A)$, so BRA is false. Hence by Lemma 1, ARB holds. Thus ARB holds for every $A \in X$. $\square$

For the rest of the proof, let B as in Lemma 2 be fixed. For each $\alpha < \omega_1$, $f_B(\alpha) < \omega_{\alpha+1}$ so $\{\beta : \beta < f_B(\alpha)\}$ has cardinality at most $\aleph_\alpha$. Hence there is a one-to-one function g_α mapping $\{\beta : \beta < f_B(\alpha)\}$ into $\{\beta : \beta < \omega_\alpha\}$.

Now suppose ARB, and let $S_A = \{\alpha : f_A(\alpha) < f_B(\alpha)\}$. Then for all $\alpha \in S_A$, $g_\alpha(f_A(\alpha)) < \omega_\alpha$. Since S_A is stationary we may apply Theorem 6 to obtain a stationary set $T_A \subseteq S_A$ and an ordinal $\gamma_A < \omega_1$ so that for all $\alpha \in T_A$, $g_\alpha(f_A(\alpha)) < \omega_{\gamma_A}$. The total number of such pairs (T_A, γ_A) is clearly at most $2^{\aleph_1} \cdot \aleph_1 = \aleph_2 \cdot \aleph_1 = \aleph_2$. Since $\aleph_{\omega_1+1}$ is regular, it follows that for some pair (T, γ), $\{A : ARB, T_A = T$, and $\gamma_A = \gamma\}$ has cardinality at least $\aleph_{\omega_1+1}$.

Of course, if X and Z are any sets of cardinality κ and λ, respectively, then the number of functions from X into Z is λ^κ. Thus the number of functions mapping T into γ is

$$\aleph_\gamma^{\aleph_1} = \max(\aleph_\gamma^{\aleph_\gamma}, \aleph_1^{\aleph_1}) = \max(2^{\aleph_\gamma}, 2^{\aleph_1}) = \max(\aleph_{\gamma+1}, \aleph_2) < \aleph_{\omega_1}.$$

It follows that there exist A_1 and A_2 such that $A_1 \neq A_2$,

$$A_1RB, A_2RB, T_{A_1} = T_{A_2} = T, \gamma_{A_1} = \gamma_{A_2} = \gamma,$$

and moreover

$$g_\alpha(f_{A_1}(\alpha)) = g_\alpha(f_{A_2}(\alpha)) \quad \text{for all} \quad \alpha \in T.$$

But now since each g_α is one-to-one, we must have $f_{A_1}(\alpha) = f_{A_2}(\alpha)$ for all $\alpha \in T$. That is to say, $\{\alpha : f_{A_1}(\alpha) = f_{A_2}(\alpha)\}$ is unbounded, contradicting the fact that $A_1 \neq A_2$. This proves the theorem. □

5. Generalizations. It is not difficult to see that Theorem 7 can be improved in several ways. For instance, in the hypothesis of Theorem 7 it is not necessary to know that $2^{\aleph_\alpha} = \aleph_{\alpha+1}$ for *every* $\alpha < \omega_1$. It will suffice to know only that $\{\alpha < \omega_1 : 2^{\aleph_\alpha} = \aleph_{\alpha+1}\}$ is a stationary set. Therefore, intuitively, if the generalized continuum hypothesis holds "fairly often" below $\aleph_{\omega_1}$, then it holds at $\aleph_{\omega_1}$.

Nor is there any reason to restrict the theorem to $\aleph_{\omega_1}$. The proof will work perfectly well for any cardinal $\aleph_\beta$ for which there is a monotonically increasing function h with domain ω_1 which satisfies $\aleph_\beta = \sup\{\aleph_{h(\alpha)} : \alpha < \omega_1\}$.

Since a treatment similar to the one in section 3 can be given for closed unbounded and stationary sets with respect to any regular uncountable cardinal in place of $\aleph_1$, appropriate generalizations of Theorem 7 hold also. On the other hand, since there is no satisfactory notion of closed unbounded or stationary subsets of ω, the proof apparently cannot be modified to show that if $2^{\aleph_n} = \aleph_{n+1}$ for all $n \in \omega$, then $2^{\aleph_\omega} = \aleph_{\omega+1}$.

There is no need to state Theorem 7 only for the generalized continuum hypothesis either. By complicating the proof only a little, it can be shown that if $2^{\aleph_\alpha} \leqq \aleph_{\alpha+2}$ for all $\alpha < \omega_1$, then $2^{\aleph_{\omega_1}} \leqq \aleph_{\omega_1+2}$. In fact, if β is any ordinal less than ω_1 and $2^{\aleph_\alpha} \leqq \aleph_{\alpha+\beta}$ for all $\alpha < \omega_1$, then $2^{\aleph_{\omega_1}} \leqq \aleph_{\omega_1+\beta}$.

Silver's original theorem **[11]** combined all these remarks.

But even further generalizations are possible, as Galvin and Hajnal **[7]** recently showed. One special case of their theorem (which is difficult to state in general) asserts the following:

Suppose only that $2^{\aleph_\alpha} < \aleph_{\omega_1}$ for all $\alpha < \omega_1$. (In this case, $\aleph_{\omega_1}$ is called a *strong limit cardinal.*) Then

$$2^{\aleph_{\omega_1}} < \aleph_{\omega_{\omega_1}}.$$

1st author's work partially supported by NSF grant No. GP-38026.
2nd author's work partially supported by NSF grant No. GP-43841.

References

1. P. S. Alexandroff and P. Urysohn, Mémoire sur les espaces topologiques compacts, Verh. Akad. Wiss. Amsterdam (1) 45, Nr. 1, 1–96.

2. J. E. Baumgartner and K. Prikry, On a theorem of Silver, Discrete Math., 14 (1976) 17–21.

3. G. Bloch, Sur les ensembles stationnaires de nombres ordinaux et les suites distinguées de fonctions régressives, C. R. Acad. Sci. Paris, 236 (1953), 265–267.

4. P. J. Cohen, The independence of the continuum hypothesis, Proc. Nat. Acad. Sci. U.S.A., 50 (1963) 1143–1148 and 51 (1964) 105–110.

5. W. B. Easton, Powers of regular cardinals, Annals Math. Logic, 1 (1970) 139–178.

6. G. Fodor, Eine Bemerkung zur Theorie der regressiven Funktionen, Acta Sci. Math., 17 (1956) 139–142.

7. F. Galvin and A. Hajnal, Inequalities for cardinal powers, Ann. Math., 101 (1975) 491–498.

8. K. Gödel, Consistency proof for the generalized continuum hypothesis, Proc. Nat. Acad. Sci. U.S.A., 25 (1939) 220–224.

9. J. König, Zum Kontinuum Problem, Math. Ann., 60 (1905) 177–180.

10. M. Magidor, On the singular cardinals problem, preprint.

11. J. Silver, On the singular cardinals problem, Proc. 1974 Int'l Congress of Mathematicians.

12. R. Solovay, $2^{\aleph_0}$ can be anything it ought to be, Proc. 1963 Symposium on the Theory of Models, North-Holland, 1965, 435.

DEPARTMENT OF MATHEMATICS, DARTMOUTH COLLEGE, HANOVER, NH 03755.
INSTITUTE FOR ADVANCED STUDY, PRINCETON, NJ 08540.

MATHEMATICAL NOTES

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Laboratoire Calcul des Probabilités, Université de Paris, T.56, 4 Place Jussieu, 75–230 Paris, France.

COMPLEMENTARY SYSTEMS OF INTEGERS

AVIEZRI S. FRAENKEL

The nonempty sequences $S_1, \ldots, S_m$ of positive integers are called *complementary* or a *complementary system*, if every positive integer occurs exactly once in exactly one of the sequences. We shall characterize two broad classes of complementary systems.

1. A simple characterization of all homogeneous complementary systems.

THEOREM 1. *Let* $m > 1$, $\alpha_1, \ldots, \alpha_m$ *positive. Then*

$$\{[n\alpha_i]: i = 1, \ldots, m;\ n = 1, 2, \ldots\} \tag{1}$$

is a complementary system if and only if (i) $m = 2$, (ii) $\alpha_1^{-1} + \alpha_2^{-1} = 1$, (iii) α_1 *irrational.*

Notes. There is another but trivial case, namely $m = \alpha_1 = 1$, which was excluded above. The symbol [ ] denotes the greatest integer function. The fact that if (1) is a complementary system then $m = 2$ was proved by Uspensky **[6]**. See also Skolem **[5]** and Graham **[2]**. The fact that for $m = 2$ (1) is complementary if and only if (ii) and (iii) hold is well-known. See e.g., **[1, 4, 7]**. Our purpose is to give a simple proof of the combined result, which relates the full story about homogeneous complementary systems.

Proof. Suppose that the system (1) is complementary. Since 1 is represented in (1), we have $[\alpha_1] = 1$, say. We first show that $[n\alpha_1]$, $[n\beta - \varepsilon]$ $(n = 1, 2, \ldots)$ are complementary, where

$$\beta = \frac{\alpha_1}{\alpha_1 - 1}, \quad \varepsilon = \begin{cases} (2(a-c))^{-1} & \text{if } \alpha_1 = a/c \text{ is rational} \\ 0 & \text{otherwise.} \end{cases}$$

Let $\zeta = \{n\alpha_1, n\beta - \varepsilon : n = 1, 2, \ldots\}$. It suffices to show that if $M > 1$ is any integer, then the number of terms of ζ less than M is $N = M - 1$. Now $n\alpha_1 < M \Leftrightarrow n\alpha_1 + \delta < M$, where $\delta = (2c)^{-1}$ if $\alpha_1 = a/c$; $\delta = 0$ if α_1 is irrational. The number of $n\alpha_1 + \delta < M$ is the same as the largest n satisfying this inequality, which is $[(M - \delta)/\alpha_1]$. Similarly, the number of $n\beta - \varepsilon < M$ is $[(M + \varepsilon)/\beta]$. Thus $N = [(M - \delta)/\alpha_1] + [(M + \varepsilon)/\beta]$. Now

$$\frac{M-\delta}{\alpha_1} - 1 < \left[\frac{M-\delta}{\alpha_1}\right] < \frac{M-\delta}{\alpha_1}, \quad \frac{M+\varepsilon}{\beta} - 1 < \left[\frac{M+\varepsilon}{\beta}\right] < \frac{M+\varepsilon}{\beta}.$$

Adding gives $M - 2 < N < M$, and so $N = M - 1$ as required.

Let k be the smallest positive integer not in $[n\alpha_1]$. Then, say, $k = [\alpha_2] = [\beta - \varepsilon] \geqq 2$. Now $n\beta - 1 \leqq [n\beta - \varepsilon] < n\beta - \varepsilon$, and so

$$[(n+1)\beta - \varepsilon] - [n\beta - \varepsilon] < (n+1)\beta - \varepsilon - n\beta + 1 = \beta + 1 - \varepsilon \leqq k + 2 - \varepsilon,$$

$$[(n+1)\beta - \varepsilon] - [n\beta - \varepsilon] > (n+1)\beta - 1 - n\beta + \varepsilon = \beta - 1 + \varepsilon > k - 1 + 2\varepsilon.$$

This implies

$$k \leqq [(n+1)\beta - \varepsilon] - [n\beta - \varepsilon] = [(n+1)\alpha_2] - [n\alpha_2] \leqq k+1.$$

The difference between consecutive terms of $[n\beta - \varepsilon]$, which is the same as the difference between consecutive *missing* integers from $[n\alpha_1]$ is thus k or $k+1$, the same as the difference between consecutive integers from $[n\alpha_2]$. This shows that the jth missing integer from $[n\alpha_1]$ is $[j\alpha_2] = [j\beta - \varepsilon]$. Since this holds for all j, $\alpha_2 = \alpha_1/(\alpha_1 - 1)$, and so (i) and (ii) hold. If α_1 is rational, then also $\alpha_2 = b/d$ is rational. Then for $j = d$, $[j\alpha_2] > [j\beta - \varepsilon]$, a contradiction. Hence (iii) holds.

Conversely, suppose that (i), (ii), (iii) hold. Letting now $\zeta = \{n\alpha_1, n\alpha_2 : n = 1, 2, \ldots\}$, it follows as above that (1) is a complementary system.

2. A simple characterization of all two-way splittings of the positive integers. We give a simple proof of a result of Lambek and Moser **[3]**.

THEOREM 2. *Let $f(n)$ be a non-decreasing function of nonnegative integers defined on the positive integers,*

$$F(n) = f(n) + n, \qquad G(n) = f^*(n) + n, \tag{2}$$

where $f^*(n)$ = number of positive integers x satisfying $0 \leqq f(x) < n$. Then $F(n)$ and $G(n)$ are complementary sequences. Conversely, every two increasing complementary sequences $F(n)$, $G(n)$ decompose into the form (2), *with f non-decreasing.*

Proof. Suppose F, G satisfy (2). It evidently suffices to show that if $M > 1$ is any integer, there are exactly $N = M - 1$ terms of F and G which are less than M.

Suppose that there are exactly k terms of G less than M. Then $f^*(k) + k < M \leqq f^*(k+1) + k + 1$. That is, $f^*(k) < M - k \leqq f^*(k+1) + 1$. Now

$$f^*(k) = \{\# x : f(x) < k\} < M - k \Rightarrow f(M-k) \geqq k.$$

$$f^*(k+1) = \{\# x : f(x) < k+1\} \geqq M - k - 1 \Rightarrow f(M-k-1) < k+1.$$

Hence $F(M-k) = f(M-k) + M - k \geqq M$, $F(M-k-1) < M$. So there are exactly $M-k-1$ terms of F less than M. Hence $N = M - 1$.

Conversely, suppose that F and G are complementary. Define $f(n) = F(n) - n$. Then f is non-decreasing, and so f^* is defined. By the first part, $H(n) = f^*(n) + n$ is a complement of F. Since every set has only one complement, $G = H$.

COROLLARY. $f^{**} = f$.

REMARK. The first and last parts of the proof of Theorem 1 can be effected by Theorem 2. Since this does not shorten the proof, a self-contained argument was preferred.

References

1. S. Beatty, Problem 3173, this MONTHLY, 33 (1926) 159. Solutions, *ibid.*, 34 (1927) 159.

2. R. L. Graham, On a theorem of Uspensky, this MONTHLY, 70 (1963) 407–409.

3. J. Lambek and L. Moser, Inverse and complementary sequences of natural numbers, this MONTHLY, 61 (1954) 454–458.

4. I. Niven, Diophantine approximations, Interscience Tracts in Pure and Applied Mathematics No. 14, Interscience, Wiley, New York, 1963.

5. Th. Skolem, On certain distributions of integers in pairs with given differences, Math. Scand., 5 (1957) 57–68.

6. J. V. Uspensky, On a problem arising out of the theory of a certain game, this MONTHLY, 34 (1927) 516–521.

7. A. M. Yaglom and I. M. Yaglom, Challenging mathematical problems with elementary solutions, vol. II, translated by J. McCawley, Jr., revised and edited by B. Gordon, Holden-Day, San Francisco, 1967.

DEPARTMENT OF APPLIED MATHEMATICS, THE WEIZMANN INSTITUTE OF SCIENCE, REHOVOT, ISRAEL.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics and Statistics, The University of Calgary, Calgary, Alberta, Canada, T2N 1N4. (From July 1976 to July 1977: Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, England.)

TWO PROBLEMS ON GEOMETRIC BODIES

JAN MYCIELSKI

The best counterparts of physical bodies which we have in geometry are the bounded regular-open Jordan measurable sets in $\mathbf{R}^3$. Let us explain this terminology. Bounded means that there exists a sphere of finite radius containing the set. Regular-open means that the set consists of all interior points of its closure. Jordan measurable means that the boundary of the set is of Jordan measure 0, i.e., if we cover $\mathbf{R}^3$ with a regular lattice of closed cubes with edges of length $1/n$ and let v_n be the volume of the union of the cubes which intersect both the set and its complement, then $v_n \to 0$ as $n \to \infty$. Let $\mathbf{B}$ denote the class of all geometric bodies, i.e., bounded regular-open Jordan measurable sets in $\mathbf{R}^3$.

It is easy to prove that if $A, B \in \mathbf{B}$ then $A \cap B \in \mathbf{B}$ and if we put

$$A \vee B = \operatorname{int}(\operatorname{cl}(A \cup B))$$

and

$$A - B = A \setminus \operatorname{cl}(B),$$

where $\operatorname{int}(X)$ and $\operatorname{cl}(X)$ denote the interior and the closure of X respectively and $\setminus$ denotes the set theoretic difference, then $A \vee B \in \mathbf{B}$ and $A - B \in \mathbf{B}$.

Two sets $A, B \in \mathbf{B}$ are called equivalent by finite decomposition, in symbols $A \cong B$, if there exist a positive integer n, disjoint sets $A_1, \ldots, A_n$ and disjoint sets $B_1, \ldots, B_n$ all in $\mathbf{B}$ and some orientation and distance preserving transformations $T_1, \ldots, T_n$ of $\mathbf{R}^3$ such that

$$A = A_1 \vee \cdots \vee A_n, \qquad B = B_1 \vee \cdots \vee B_n$$

and

$$T_i(A_i) = B_i \quad \text{for} \quad i = 1, \ldots, n.$$

Notice the following easy facts:

(i) $\cong$ is an equivalence relation in $\mathbf{B}$.
(ii) If $A \cong B$ then $\operatorname{vol}(A) = \operatorname{vol}(B)$.
(iii) If $A \cong B$ and $A \subseteq B$ then $A = B$.

For any $A \in \mathbf{B}$ let $|A|$ be the type of A, i.e., $|A| = \{B \in \mathbf{B}: B \cong A\}$. For any $A, B \in \mathbf{B}$ we put

$$|A| + |B| = |A \vee T(B)|,$$

where T is any orientation and distance preserving transformation of $\mathbf{R}^3$ such that $A \cap T(B) = \varnothing$. Also

$$n|A| = |A| + |A| + \cdots + |A|$$

with n summands on the right. Finally,

$$|A| \leqq |B|$$

means that there exists $B_0 \subseteq B$, $B_0 \in \mathbf{B}$ such that $B_0 \cong A$. By (iii) we have

(iv) If $|A| \leqq |B|$ and $|B| \leqq |A|$ then $|A| = |B|$.

PROBLEM 1. Given a positive integer n and $A, B \in \mathbf{B}$, does

$$n|A| = n|B| \Rightarrow |A| = |B|, \tag{1}$$

$$n|A| \leqq n|B| \Rightarrow |A| \leqq |B|? \tag{2}$$

In fact (1) is open already for $n = 2$. The converse of (ii) is also an open question. By (iv) we have (2) $\Rightarrow$ (1).

A number of algebraic facts about our types follow from the theory of refinement algebras (R.A.) introduced by Tarski **[5]**, Definition 11.26. Our types with the operation $+$ constitute an R.A. and the Theorem 11.28 of **[5]** applies to them. On the other hand, as pointed out in **[5]**, many algebraic problems in the style of (1) and (2) remain open.

Now let us consider the class $\mathbf{B}^*$ of all bounded regular-open sets in $\mathbf{R}^3$. Thus $\mathbf{B} \subseteq \mathbf{B}^*$. For all A, $B \in \mathbf{B}^*$ we define $A \vee B$, $A - B$, $A \cong B$, $|A|$, $|A| + |B|$ and $|A| \leqq |B|$ in the same way as above except that we substitute everywhere $\mathbf{B}^*$ for $\mathbf{B}$. Then (i) holds and $\mathbf{B}^*$ has the following closure property (not true for $\mathbf{B}$)

(v) If $\mathbf{A} \subseteq \mathbf{B}^*$ then $\operatorname{int}(\bigcap \mathbf{A}) \in \mathbf{B}^*$.

This is fortunate, since, although (ii) fails for $\mathbf{B}^*$, nevertheless (v) yields (iv) (by an easy modification of the standard proof of the Cantor-Bernstein theorem, see e.g. **[3]**, p. 325). (v) also implies that our algebra of types of sets in $\mathbf{B}^*$ is again an R.A. (by **[5]**, Theorem 16.8). Problem 1 is open again. Also we do not know if (iii) is true for $\mathbf{B}^*$ and this will be stated below as Problem 2. First let us prove the following

THEOREM. *The following propositions are equivalent to each other*

(a) *There exist $A, B \in \mathbf{B}^*$ such that $A \cong B$, $A \subseteq B$ and $A \neq B$.*

(b) *For every non empty set P, $Q \in \mathbf{B}^3$ we have $P \cong Q$.*

(c) *If C is the open unit cube in $\mathbf{R}^3$ then $2|C| = |C|$.*

Proof. Notice that (b) $\Rightarrow$ (c) $\Rightarrow$ (a) are obvious. It remains to show that (a) $\Rightarrow$ (b). Let $P, Q \in \mathbf{B}^*$, $P \neq \varnothing \neq Q$ and A and B be given by (a). Then there exists a similarity S and a translation T such that $S(A) \subseteq P$ and $TS(A) \subseteq Q$. Clearly $S(A) \cong S(B)$, $TS(A) \cong TS(B)$, and there exists a cube $C_0 \subseteq S(B) \backslash S(A)$ and $T(C_0) \subseteq TS(B) \backslash TS(A)$. By (iv) for $\mathbf{B}^*$ we have $|S(A)| + |C_0| = |S(A)|$. Hence for every positive integer n we have $|S(A)| + n|C_0| = |S(A)|$. Also, for n large enough, $n|C_0| \geqq |P|$ and $n|C_0| \geqq |Q|$. Hence

$$|S(A)| \leqq |P| \leqq |S(A)| + n|C_0| = |S(A)|$$

and, by (iv), $|P| = |S(A)|$. In the same way we show that $|Q| = |TS(A)| = |S(A)|$ and (b) follows.

Q.E.D.

PROBLEM 2. Do (a), (b) and (c) fail? In other words, do there exist any types of sets in $\mathbf{B}^*$ different from $|\varnothing|$ and from $|C|$?

Notice that there is no such problem for $\mathbf{B}$. That is, by (ii) and (iii), the analogs of (a), (b) and (c) for $\mathbf{B}$ are all false.

Let us add that for the plane $\mathbf{R}^2$ Problem 1 is open, but Problem 2 has been solved and, as expected, the answer is yes. This is a Theorem of E. Marczewski (see **[3]**) who realized that this follows from a more general theorem of Banach. Strictly speaking Marczewski's theorem is the following.

There exists a finitely additive non-negative measure m over the algebra of all bounded sets in $\mathbf{R}^2$ which equals the Jordan measure for all Jordan measurable sets, is invariant under all distance preserving transformations of $\mathbf{R}^2$ and vanishes for all meager sets. (Moreover, m can be such that $m(aX) = |a|^2 m(X)$, where $aX = \{(ax, ay): (x, y) \in X\}$.) Clearly this measure m, unlike the Lebesgue measure, has the property that if $A, B \in \mathbf{B}^*$ and $A \cap B = \varnothing$ then

$$m(A \vee B) = m(A) + m(B).$$

Hence $A \cong B \Rightarrow m(A) = m(B)$ and if C is the unit square then $2|C| \neq |C|$ follows. But if instead of $\mathbf{R}^2$ we take the spherical surface $\mathbf{S}^2 = \{(x, y, z): x^2 + y^2 + z^2 = 1\}$ then Problem 2 reappears (here we can take $C = \{(x, y, z) \in \mathbf{S}^2: z < 0\}$). For $\mathbf{S}^1 = \{(x, y): x^2 + y^2 = 1\}$, Marczewski's theorem is again valid and Problem 2 is settled. These solutions rest on the fact that the groups of distance preserving transformations of $\mathbf{R}^2$ and $\mathbf{S}^1$ are solvable; for $\mathbf{R}^3$ and $\mathbf{S}^2$ they are not solvable and by a theorem of Hausdorff, later refined by Banach and Tarski (the existence of paradoxical decompositions of the sphere, see [3]), Marczewski's theorem fails for those spaces. Still Problem 2 may well have a positive solution for all $\mathbf{R}^n$ and all $\mathbf{S}^n$.

For related facts and references see [**1**, **2**, **3**, **4**, **5**].

References

1. R. E. Bradford, Cardinal addition and the axiom of choice, Ann. Math. Logic, 3 (1971) 111–196.

2. A. P. Morse, Squares are normal, Fund. Math., 36 (1949) 35–39.

3. J. Mycielski, Commentaries to some papers of Banach, in S. Banach, Oeuvres, Vol. I, Warsaw 1967, pp. 318–327.

4. A. Tarski, Algebraische Fassung des Massproblems, Fund. Math., 31 (1938) 47–66.

5. ———, Cardinal Algebras, Oxford University Press, New York, 1949.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COLORADO, BOULDER, CO 80309.

CLASSROOM NOTES

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Laboratoire Calcul des Probabilités, Université de Paris, T.56, 4 Place Jussieu, 75-230 Paris, France.

A COUNTEREXAMPLE IN CONNECTION WITH EGOROV'S THEOREM

WOLFGANG WALTER

Let (f_n), be a sequence of real-valued, measurable functions, defined on $D = [0, 1]$, which converges pointwise, and let

$$f(x) := \lim_{n \to \infty} f_n(x) \quad \text{for} \quad x \in D. \tag{1}$$

Egorov's theorem [**1**; p. 88] [**2**; p. 37] states that the sequence converges "almost uniformly," i.e., that for any given $\varepsilon > 0$ there exists a measurable set $N_\varepsilon \subset D$ with Lebesgue measure $\lambda(N_\varepsilon) < \varepsilon$ such that the convergence is uniform on $D \dot{-} N_\varepsilon$. [The theorem holds on every finite measure space D.]

Professor Ostrowski recently* raised the question if the corresponding statement is true if, instead

* (at the 13th International Symposium on Functional Equations in Oberwolfach/Germany, July 6–12, 1975)

of a sequence (f_n), we have a family (f_t) of measurable functions, say, for real $t \geqq t_0$, such that the following limit

$$f(x) := \lim_{t \to \infty} f_t(x), \qquad x \in D, \tag{2}$$

exists.

At first sight, an affirmative answer is easily at hand. For, the sequence (g_n) defined by

$$g_n(x) := \sup\{|f_t(x) - f(x)| : n \leqq t < \infty\} \tag{3}$$

converges to zero as $n \to \infty$. According to Egorov's theorem, it converges almost uniformly, which in turn implies that (2) also converges almost uniformly. But this reasoning breaks down, since, in general, the g_n are not measurable.

We give a counterexample of a family of functions (f_t) defined for $t \geqq 2$, measurable in x (for fixed t) and in t (for fixed x). It has the property that

$$\lim_{t \to \infty} f_t(x) = 0 \qquad \text{for} \quad x \in D, \tag{4}$$

but it does not converge almost uniformly.

The construction uses a nonmeasurable set $A \subset [0, \frac{1}{2})$ with the property that the sets $r + A$ and $s + A$ are disjoint for any two *rational* numbers $r, s (r \neq s)$. It is well known that such a set A exists [**2**; p. 22]. The sets $A_n := A + (1/n)$ $(n = 2, 3, \ldots)$ are pairwise disjoint nonmeasurable subsets of D, and $\lambda^*(A) = \lambda^*(A_n) =: \alpha > 0$, where λ^* denotes outer Lebesgue measure.

We define

$$f_{n+x}(x) = 1 \qquad \text{for} \quad x \in A_n \qquad (n = 2, 3, \ldots) \tag{5}$$

and $f_t(x) = 0$ at all other points (x, t) in the strip $D \times [2, \infty)$. In other terms, let Q_n denote the square $D \times [n, n+1]$ in the (x, t)-plane and $D_n := \{(x, n + x) : x \in D\}$ a diagonal of Q_n $(n \geqq 2)$. We set $f_t(x) = 1$ at all points $(x, t) \in D_n$ such that $x \in A_n$ and $f_t(x) = 0$ at all other points in Q_n $(n = 2, 3, \ldots)$.

This function $f_t(x)$ has the required properties. Obviously, for fixed t, $f_t(x)$ is zero in D with the possible exception of one point x where $f_t(x) = 1$. Likewise, since the A_n are pairwise disjoint, $f_t(x)$ is, for fixed x, zero in $[2, \infty)$ with the possible exception of one value of t. From here, the statements about measurability and the relation (4) follow. We have

$$g_n(x) := \sup\{f_t(x) : t \geqq n\} = 1 \quad \text{for} \quad x \in B_n := A_n \cup A_{n+1} \cup \ldots \tag{6}$$

and $g_n(x) = 0$ otherwise.

Let us assume that $f_t(x) \to 0$ $(t \to \infty)$ uniformly on $D - N$. Then to $\varepsilon = \frac{1}{2}$ there exists an integer $p > 0$ such that $f_t(x) < \frac{1}{2}$ for $x \in D - N$, $t \geqq p$. Because of (6) we have $N \supset B_p \supset A_p$, hence $\lambda^*(N) \geqq \lambda^*(A_p) = \alpha > 0$. This inequality shows that (f_t) is not almost uniformly convergent.

Final remarks. (a) The function $f_t(x)$ constructed above is also measurable in $(x, t) \in D \times [2, \infty)$. (b) If the functions g_n defined by (3) are measurable, then we have almost uniform convergence in (2). For example, this is the case if $f_t(x)$ is continuous in t for fixed x (because in (3) one may assume that t is rational).

References

1. P. R. Halmos, Measure Theory, Van Nostrand, Princeton, N. J., 1950.
2. J. C. Oxtoby, Measure and Category, Springer Verlag, New York, Heidelberg, Berlin 1971.

UNIVERSITÄT KARLSRUHE, MATHEMATISCHES INSTITUT I, ENGLERSTRASSE 2, POSTFACH 6380, D75 KARLSRUHE 1, WEST GERMANY.

EDITED BY PAUL T. MIELKE AND SHIRLEY HILL

Material for this Department should be sent to Paul T. Mielke, Department of Mathematics, Wabash College, Crawfordsville, IN 47933, or to Shirley Hill, Department of Mathematics, University of Missouri, Kansas City, MO 64110.

THE PERSONALIZED SYSTEM OF INSTRUCTION IN INTRODUCTORY CALCULUS

KENNETH F. KLOPFENSTEIN

Introduction. Keller [1] first described the instructional system which has come to be called the Personalized System of Instruction or, simply, PSI. During the 1973–74 academic year PSI was used in the first term of the calculus sequence offered for physical science and engineering students at Colorado State University. Evaluative data was systematically collected during the Winter Quarter. Some of this data and some conclusions drawn from it are presented here. While data from Fall and Spring Quarters is less complete, it indicates that similar results were obtained. Experience with this course suggests that implementation of PSI in calculus courses may present difficulties which have not been encountered in more elementary mathematics courses and in other disciplines.

The Personalized System of Instruction. The name PSI refers to a specific system of instruction which, according to Keller [1], is distinguished by the following features: (a) It is self paced. (b) Mastery is required for advancement to new material. (c) Course content is communicated primarily through the written word. (d) Peers are employed as course assistants to allow individualized scoring of repeated tests. While systems which do not embody these features may be effective and educationally sound, by definition, they are not PSI.

In a PSI course, the content is divided into short units which are to be studied and mastered in sequence. A printed study guide which specifies the objectives of the unit, suggests resources for accomplishing the objectives, and sometimes provides supplementary materials is prepared for each unit. When a student feels prepared, he or she takes a short quiz to verify his or her mastery of that material. Immediately when the quiz is completed, a course assistant (usually called a proctor or tutor) reads the quiz with the student. If the student has achieved a perfect score, he or she proceeds to the next unit. If not, he or she is given advice for further study and takes a different quiz on the same material. Unit quizzes are repeated until mastery is demonstrated. Repeating quizzes in no way affects the student's grade. In most PSI courses, motivational lectures that are open only to students who have completed the prerequisite units are given periodically. For a more detailed description of a PSI course, see Keller [1].

An indication of the diversity of disciplines and institutions where PSI has been applied is given by the *PSI Newsletter* [4]. Kulik, Kulik, and Carmichael [2] have reviewed the evaluative research on PSI courses. Their conclusions show that the effectiveness of the method, as measured by student attitudes and achievement, should no longer be a matter of serious debate. The question in applying PSI to teach calculus, then, should not be "can PSI be used effectively to teach calculus?", but, instead, "how should PSI be implemented in a calculus course in order that it be as effective as experience indicates that it can be?"

Implementation. Every effort was made to implement PSI in Calculus I as nearly as possible as Keller [1] describes it. However, no lectures were given because it was believed that the time could be more profitably used to prepare better study guides. The class met for 50 minutes daily for 10 weeks. Unit quizzes were given only during these times. The content consisted of the first five chapters of Thomas [5]. This material was divided into twenty units, including four review units. Students were

obliged to complete all of the units to receive a grade. Unit twenty was a comprehensive review and sampled the entire course. The unit twenty quiz could be repeated and emphasized, but did not require mastery. The student's course grade was determined by his or her best score on the unit twenty quiz. For this reason, the unit twenty quiz was scored by the instructor. A score below 80% earned a C, a score between 80% and 90% earned a B, and a score above 90% earned an A. Students who did not complete unit twenty by the end of the Winter Quarter were assigned a grade of incomplete. These students were allowed one additional quarter to complete the course, either by continuing in the PSI mode or by enrolling in another section of the course. Students who continued in the PSI mode but did not complete the course in the time allowed were required to take the Spring Quarter departmental final examination which was used to determine their final grades.

Comparison sections. The performance and attitudes of the PSI class were compared with those of a class taught in a lecture-discussion style by a professor who is recognized as an effective teacher. These two groups were found to be essentially the same with respect to sex, class in school, SAT verbal score, ACT verbal score, high school grade point average, high school rank in class, grades in previous university mathematics courses, and university grade point average. However, students in the PSI group were more likely to have declared majors in engineering curricula, and had slightly higher mean scores on both the SAT and ACT mathematical aptitude tests. A difference in mathematical aptitude was also indicated by the fact that a larger proportion of the PSI students had obtained credit in College Algebra by examination.

Withdrawal and completion rates. Fifty-three students were enrolled in the PSI section. Table 1 shows how these students finally disposed of the course.

TABLE 1: *Student's Disposition of Calculus I, PSI*

Initial Enrollment	53	
Drops		
— Withdrawals (Formal & Informal)	12	38%
— To Lecture Next Term	8	
Net Enrollment	33	
Completed in One Quarter	17	52%
Completed in Two Quarters	8	24%
Took Final, Spring Quarter	8	24%

According to available records, about 17% of the students in all other sections of the course withdrew or received a grade of F.

Table 2 shows the distribution of grades for the 33 students who completed the course.

TABLE 2: *Final Grades, Calculus I, PSI*

Completed in One term	Completed in Two terms	Took Final Examination	Total	
A—15	A—1	A—1	A—17	(52%)
B—2	B—1	B—2	B—5	(15%)
	C—6	C—2	C—8	(24%)
		D—1	D—1	(3%)
		F—2	F—2	(6%)

Post test. A twenty item examination covering Calculus I was administered early in Spring Quarter to all students in Calculus II. In Table 3 the scores of students from the PSI class are compared with those of students who received an A or B grade in the comparison section and with A

and B students from all traditional sections. (The mean examination score for each group is denoted by $\bar{x}$.)

TABLE 3: *Comparison of Post Test Scores*

	PSI		Comparison Section		All Tradit. Sect.	
	A Grade	All	A Grade	A or B Grades	A Grade	A or B Grades
n	10	14	3	10	15	39
$\bar{x}$	10.6	10.9	9.3	8.5	9.9	8.6
s.d.	3.9	3.4	1.7	1.4	1.8	2.3

Clearly the students from the PSI section performed as well as the A and B students from the traditional sections of the course. The mean scores from the PSI group were higher but not significantly so. Recall that there were indications that the PSI group had higher mathematics aptitude than the comparison group, and this might account for the higher mean scores.

Comparison of attitudes. During the seventh week of the quarter an adaption of the Neidt Attitude Scale **[3]** was administered to a sample of 28 students in the PSI section and 27 students in the comparison section. The Neidt Attitude Scale yields three scores relating to attitudes toward method of instruction, course content, and the extent to which the course has met the student's expectations, as well as a total score. The scale has shown satisfactory reliability in a variety of situations. Table 4 summarizes the results. High scores reflect favorable attitudes.

TABLE 4: *Scores from the Neidt Attitude Scale*

		$\bar{x}$	s.d.	Range of Scores	Possible Range	Midpoint of scale
Method	PSI	19.42	4.7	10–27	6–30	18
	Comparison	21.62	7.4	11–29		
Content	PSI	18.32	2.8	13–25	5–25	15
	Comparison	17.22	4.2	7–25		
Expectation	PSI	15.50	4.1	8–23	5–25	15
	Comparison	17.11	3.2	8–22		
Total	PSI	53.25	9.9	34–71	16–80	48
	Comparison	56.37	11.2	29–76		

The attitudes of the students in the comparison section were slightly more favorable than those of students in the PSI section, but none of the mean differences were statistically significant. The fact that the mean scores differed slightly from the midpoint of the scale indicates that, on the whole, both groups were quite neutral in their attitudes. However, in each group there were students who were very favorable and others who were very unfavorable.

During the sixth and seventh weeks of the quarter, eighteen students from the PSI group were interviewed by telephone. Fifteen of the students who were interviewed were actively involved in the course, while three had informally withdrawn and were planning to complete the course in a traditional section during the next quarter. Of the fifteen active students, eleven were on or ahead of schedule and four were behind schedule.

The following statements made in the interviews summarize the student's general feelings about the method of instruction. The numbers in parentheses indicate how many respondents made the statement.

Positive Comments:

I like the method, I learn better.

I like it, but the course needs more work. (2)

I like it and would like it for Calculus II.

I like the idea of having two quarters to complete the course. (2)

Neutral Comments:

I am skeptical.

It went well at first but has become awfully time demanding.

Mixed feelings. There is a lot of pressure to keep up.

It is OK, but there is not enough help.

Pretty good, but hard. Proctors are a help.

Makes you learn, hard at start.

It is good if you want to know calculus.

It is OK if you have the motivation.

Negative Comments:

I don't like it. You have to teach yourself and I need the lecture. (2)

I don't like the method. There is not enough guidance.

There are too many picky details on the tests.

The overall impression from the telephone interviews was again one of neutrality.

Conclusions. Even though established procedures for implementing and managing PSI courses were carefully followed, the PSI calculus course was much less successful than the literature indicates it could be. Withdrawal and failure rates were abnormally high, and an unacceptably large number of students required two terms to complete the course. However, those who completed the course in one term did earn the expected grades of A or B, and their performance on a post-test verified a high level of competency. No significant difference in attitude between students in the two sections was detected.

The problem is to determine how this implementation of PSI should be modified to achieve the successes reported in the literature. There are at least three factors which may have contributed to the limited success of the system in this situation.

First, the amount of content in the PSI course probably exceeded that expected of even the A student in traditional sections. The material *covered* in the two methods was nearly the same. However, with PSI every student was *required* to learn *all* of that material while with traditional methods only 15% of the class (the "superior" students) were *expected* to learn about 85% of the material covered. This observation is not intended as a criticism of either mode of instruction. It is merely a difference that must be reckoned with in the selection of content and the formulation of objectives for a PSI calculus course.

Informal observations of errors and conceptual difficulties that students encountered suggest that many students who begin the study of calculus have not acquired the prerequisite algebraic, geometric, and analytic skills and concepts that are needed to master the subject. This deficiency could be remedied by integrating a review of the necessary prerequisites into each unit.

Finally, many beginning calculus students have little confidence in their ability to do mathematics and are convinced that mathematics is impossibly difficult. Because of this aversion to mathematics, students in a PSI calculus course put off studying the material. The course can be structured to respond to this problem by designing the early units to insure that students will be quickly successful in completing them. Instituting biweekly motivational lectures open only to those students who are progressing on schedule should also help solve this problem.

Assuming that PSI is a sound instructional system, iteration of the course design is the natural next step. Although the first iteration was less than successful, it served to identify some needed modifications. These modifications can be incorporated into a second iteration. The second iteration will no doubt reveal changes that should be incorporated into a third iteration, and so on. One of the important advantages of PSI is that it can provide all of the information needed to systematically improve the course by iterating the course design. Perhaps the experience reported here can help to improve the first iteration of PSI in the calculus at other colleges and universities.

Acknowledgements. This project was undertaken with the encouragement of Professors Fred S. Keller and Billy V. Koen. Several discussions with them during the planning of the course were very helpful. The Colorado State University Office of Instructional Development provided funds for duplicating course materials and provided the assistance of Professor Douglas Sjogren, who conducted the telephone interviews and processed the evaluative data.

References

1. F. S. Keller, Goodbye teacher..., J. of App. Beh. Anal., 1 (1968) 79–89.

2. J. A. Kulik, C. L. Kulik, and K. Carmichael, The Keller plan in science teaching, Science, 183 (1974) 379–383.

3. C. O. Neidt and D. D. Sjogren, Changes in student attitudes during a course in relation to instructional media, AV Communication Review, 16 (1968) 268–279.

4. PSI Newsletter, The Center for Personalized Instruction, Georgetown University, Washington, D. C.

5. G. B. Thomas, Jr., Calculus and Analytic Geometry, Alternate Edition, Addison-Wesley, Reading, Mass., 1972.

DEPARTMENT OF MATHEMATICS, COLORADO STATE UNIVERSITY, FORT COLLINS, CO 80523.

SELF-PACED MATHEMATICS INSTRUCTION: A STATISTICAL COMPARISON WITH TRADITIONAL TEACHING

ADA PELUSO AND A. J. BARANCHIK

1. Introduction and summary. Two modes of self-pacing instruction, one of which is based on the Keller plan, are compared with instruction in the traditional classroom manner. In an introductory basic concepts of mathematics course, we find the two self-pacing methods to be superior to traditional teaching, both in terms of final exam performance and success in subsequent mathematics courses. In particular, one of the self-pacing methods appears to raise weak (as measured by a pretest) students' performance up to the level of that of strong students. These methods were also compared in teaching beginning calculus, with more mixed results. In section 2 the modes of instruction are described. Sections 3 and 4 contain statistical comparisons of the teaching methods in the introductory course and in calculus, respectively. These experiments took place at Hunter College of the City University of New York (CUNY) in the fall of 1972. A more detailed description of the Mathematics Learning Center at Hunter College and of the modes of instruction is given in section 5.

2. The three modes of instruction. Basic Structures of Mathematics, a one-semester course intended for non-science students, has a syllabus consisting of the elements of logic, relations and operations, elementary number theory, set theory, the rudiments of real numbers, some analytic geometry, and an introduction to functions and their graphs. 617 students registered for this course in the fall of 1972, each student receiving one of three modes of instruction:

Classroom (C): Regular classroom instruction, consisting of 3 lecture hours per week, with one instructor for approximately 35 students.

Learning Center (LC): Self-pacing instruction utilizing tape-slide shows, with one instructor and one tutor available for each set of approximately 35 students.

Learning Center-Keller Plan (LCK): Similar to LC but with a Keller plan modification in which each group of 10 to 12 students was identified with a student tutor.

These three modes of instruction were also compared for teaching a first semester course in the calculus, which results are reported in section 4. In both courses the final exam was prepared from questions solicited from all the instructors of the course in order to assure input from all three teaching modes.

TABLE 1. *Raw Data for Basic Structures Course*

	Enrolled	Passed	Failed	Dropped	Incomplete	Number Taking Final	Average Final Score
C	241	188	22	31	0	203	63.91
LC	190	138	8	34	10	145	77.63
LCK	186	127	24	23	12	142	74.20
Totals	617	453	54	88	22	490	70.95

A one-way analysis of variance (ANOVA) showed significant difference in final exam scores ($F_{2,487} = 33.292$) among the three instruction modes, but no significant difference ($F_{1,285} = 1.656$) between the two self-pacing (LC, LCK) methods. However, this crude analysis does not allow for possible differences in inherent ability of the students in the various types of classes; hence, the more detailed analysis of section 3.

3. A comparison of the three modes of teaching. In order to avoid finding differences among teaching methods which might, in fact, be due to differing levels of preparation on the part of students, analysis was restricted to those students who had been pretested in arithmetic computation. Thus, final exam performances of comparable students can be compared (Tables 2 and 3). Moreover, performances of students in their next mathematics course (traditionally taught) were also compared (Table 4) and these results were consistent with the differences found in final exam performances. Finally, in an attempt to study the value of the pretest as a predictor (linear regression) of final exam performance (Table 5), we were led to discover that LC had a leveling effect on student performance — a student's final exam score was virtually independent of his pretest score. But this leveling was efficacious, i.e., it was a "leveling up," the weak students being brought up to the level of stronger students. This property of LC was somewhat supported by data for calculus students (Table 9, section 4) as well.

TABLE 2. *Final Exam Performance of Pretested Students*

	No. Pretested	No. Taking Final	Avg. Final Score	Pretest Range (60 questions)	Pretest Median
C	86	78	68.71	14–58	37.5
LC	54	49	81.53	21–55	41
LCK	45	39	76.44	14–55	40

Comparing Table 2 with Table 1 it is evident that pretested students (entering freshmen) as a group performed better than non-pretested students. Also, a higher proportion of traditionally taught (C) students were pretested.

TABLE 3. *F-tests for 2-Way ANOVA*

Teaching Method	$F_{2,157} = 15.9$
Pretest Performance	$F_{2,157} = 25.0$
Interaction	$F_{4,157} = 2.72$

From Table 3, *the apparent superiority of LC and LCK over C as seen in Table 2 is strongly supported.* An F-test was also performed to compare LC with LCK. $F_{2,82} = 3.76$, barely significant at .05 level, indicates a slight superiority of LC over LCK.

Although there was no significant difference in the proportion of students who, in the following semester, chose to take a mathematics course which follows Basic Structures in the curriculum, among those who did so *those trained in the Learning Center (LC, LCK) were more likely to pass their next mathematics course* ($\chi_2^2 = 6.3$, significant at .05 level).

TABLE 4. *Follow-up Data*

	Passed Basic Structures	Continued and Passed	Continued Unsuccessfully	% Continuers
C	188	37	17	28.7
LC	141	38	5	30.5
LCK	128	35	7	32.8

(N.B. The number of passers in Table 4 differs from Table 1 because some students raised their grades of Incomplete to passing grades.)

Finally, in an attempt to predict final exam performance (y) as a function of pretest score (x), we obtained the following regressions.

TABLE 5. *Linear Regression*

	Sample Size	Regression Line ($y =$)	t-value	Significance
C	78	$28.3 + 1.1x$	6.27	$< .0005$
LC	49	$64.1 + .42x$	1.65	$> .05$
LCK	39	$28.2 + 1.2x$	5.66	$< .0005$

From this we obtain the striking conclusion that performance in C and LCK is strongly related to prior mathematical skills but that this is not so for LC. Bearing in mind the high overall final exam average in LC (see Table 2), one might suspect this lack of dependence of final upon pretest indicates that *weak* (on pretest) *students were brought up to the level of stronger students.* The actual data corroborate this conjecture: of the 9 weakest students in LC (those with pretest scores prejudged as dangerously low) 8 took the final, their final scores ranging from 74% to 95%, *all* above the average of all pretested students in the classroom (C).

4. Calculus results. In the fall of 1972 the three modes of instruction were applied to teaching 395 students a first semester course in the calculus. Comparisons were made on the basis of a common final and in terms of follow-up performance in (traditionally taught) second semester calculus. Some students (typically freshmen) were pretested with a calculus readiness test. Although results were mixed, it was most interesting to observe, even more strongly than in the Basic Structures course (see Table 5 in section 3), the "leveling" effect of the LC mode of instruction (Table 9).

TABLE 6. *Raw Data for Calculus*

	Enrolled	Passed	Failed	Dropped	Incomplete	No. Taking Final	Avg. Final Score
C	112	99	8	4	1	104	62.33
LC	142	71	37	14	20	92	60.88
LCK	141	72	27	17	25	83	61.96
Totals	395	242	72	35	46	279	61.74

The data in Table 6 may be misleading as there is evidence of an imbalance of stronger students being enrolled in the traditional classroom (C) sections. Pretested students (typically freshmen) were more likely to enter a C section, and such students — having been screened — appeared to be more likely to perform well.

TABLE 7. *Pretested Students*

	No. Pretested	Final Takers	Avg. Final Score	Pretest Range (50 questions)	Pretest Median	Pretest Avg. of Final Takers
C	79	75	63.7	21–47	34	35
LC	31	27	69.2	23–46	36	37
LCK	32	27	69.2	17–49	35	35
Totals	142	129	66.0			

TABLE 8. *F-tests for 2-Way ANOVA*

Teaching Method	$F_{2,120} = 11.6$
Pretest Performance	$F_{2,120} = 35.1$
Interaction	$F_{4,120} = 2.99$

From Table 8, among comparable students there is a significant superiority — in terms of final exam performance — of students having either of the two Learning Center modes of instruction. However, Table 7 indicates the magnitude of this difference is not so great as in the case of the Basic Structures course (Table 2), and this advantage must be weighed against the substantially lower successful-completion rate of LC and LCK as displayed in Table 6. Follow-up data were inconclusive, both in terms of likelihood of taking the subsequent calculus course and of performance in that course.

To predict final exam performance (y) as a function of pretest score (x), we obtained Table 9.

TABLE 9. *Linear Regression*

	Sample Size	Regression Line (y =)	t-value	Significance
C	75	$-8.15 + 2.045x$	4.949	$< .0005$
LC	27	$39.99 + .784x$	1.092	$> .10$
LCK	27	$2.65 + 1.878x$	3.043	$< .0005$

Again, as in Table 5 for Basic Structures, performance in LC was relatively independent of the precalculus pretest. However, in this instance 4 of the lowest (in terms of pretest) 10 students did not even take the final and, of the 6 who did, only 2 scored 60% or higher. Rather, the low t-value seems to reflect weak final exam performance at the top: the 3 top (in terms of pretest) students all had final exam scores in the 70's. Although inconclusive, this may suggest that LC is inappropriate for the very best students.

5. The mathematics learning center. With the advent of an open admissions policy at CUNY in 1970, the department of mathematics at Hunter College was determined to provide students the opportunity to complete a college-level mathematics course regardless of the students' mathematical backgrounds. Since, under open admissions, all graduates of New York City high schools were assured admission in either a senior or community college, it was expected that most freshmen would begin with Basic Structures of Mathematics, a one-semester three-credit course for non-science students. It was decided that this course should be self-paced with individualized instruction.

Traditional classroom lectures were replaced by tape-slide shows viewed in carrels, with one instructor and one student tutor present for approximately every 35 students in the class, each student progressing at his/her particular pace and taking exams when he/she felt the required material had

been mastered. This is the LC method of instruction referred to in Section 2. Each tape-slide show is a lecture on a tape cassette with a coordinated set of color slides. Preparation of these tapes and slides took place entirely within the department of mathematics.

It was anticipated that nearly 1000 students would be enrolled in the Basic Structures course that fall of 1970. With funds provided by the Alfred Sloan Foundation and a National Science Foundation Instructional Scientific Equipment Grant, 90 carrels were purchased, and each was equipped with a carousel slide projector, a tape cassette player, an eye-level projection module with a viewing screen, a headset, and a fluorescent light.

A "Study Unit in Mathematics" (SUM) was prepared for each class meeting. As each SUM provides behavioral objectives, new vocabulary and symbols, the background necessary in algebra and geometry, and a selftest with answers, students know exactly what they are expected to be able to do. Each SUM also gives students the resources available for mastering the unit and a posttest with answers for checking their understanding. Furthermore, the availability of sample exams with answers (except for the final) greatly reduces the fear of tests.

In 1971, first semester calculus joined the Basic Structures course in the Learning Center. (A self-paced precalculus course has subsequently been developed.) Calculus was offered under exactly the same setup of individualized instruction as Basic Structures.

In the LCK setup referred to in Section 2, the class met in the Learning Center under the Keller plan. Every 10 to 12 students were identified with a student tutor who assisted the students through the course, graded their exams, kept a record of their attendance and progress, and informed the instructor about the students' performance and about topics causing difficulty. The instructor of an LCK class supervised the tutors, gave occasional lectures, and assigned final grades. Extensive studies have been made of the Keller plan, and a review **[1]** of this evaluative research establishes, among other points, that students enjoy Keller courses much more than conventional courses because of self-pacing and interaction with tutors.

For the analysis, the students in Basic Structures and in first semester calculus were divided into three groups, one for each of the teaching methods. During every scheduled class hour, two teaching modes were in operation; about 35 students were taught conventionally in the classroom while another 70 students met in the Learning Center, either with the LC of the LCK setup.

For the first two years of the Learning Center, an interim grade of Y was assigned to students who needed more than one semester to finish the course. Beginning in the fall of 1972, entering freshmen at Hunter College were given a proficiency examination in mathematics computation to determine whether remediation was necessary in this area. (This led to the development of a self-paced remediation program in arithmetic and the rudiments of algebra.) The minimal proficiency requirement resulted in a smaller enrollment in Basic Structures, as well as the elimination of the Y grade; students with facility in computation were expected to complete the course within one semester. Also starting in the fall of 1972, students registering for calculus were required to take the calculus readiness test administered by the department of mathematics. On the basis of this test score, students were advised as to whether to proceed directly to calculus or to take first the three-credit precalculus course. The Y grade was no longer used in calculus either. Only students taking the courses for the first time were considered in the evaluation, and, to enable students with a Y grade to continue in the Learning Center, the classroom mode was restricted to incoming freshmen. With the mathematics computation proficiency test and the calculus readiness test it was possible in the study of the three teaching modes to compare students of comparable ability by restricting the study to those students who had been pretested.

Acknowledgments. The idea of the Mathematics Learning Center was proposed by Mary P. Dolciani and carried out under her direction and guidance. The authors wish to thank Joann Benson for organizing the data and carrying out the computations.

Reference

1. James A. Kulik, Chen-Lin Kulik, and Kevin Carmichael, The Keller plan in science teaching, Science, 183 (1974) 379–383.

DEPARTMENT OF MATHEMATICS, HUNTER COLLEGE OF CUNY, NEW YORK, N. Y. 10021.

SELF-PACED CALCULUS: A PRELIMINARY EVALUATION

RUTH REBEKKA STRUIK AND ROBERTA J. FLEXER

A study conducted at the University of Colorado in the spring of 1973 compared a traditionally taught calculus course with a self-paced calculus course. Based on performance in a common final examination, students in the self-paced course did substantially better than those in the traditional course. Prior mathematical ability and cumulative grade point average were taken into account. The data support the conclusion that self-paced instruction is more effective for many students.

The experiment — Students who enrolled in a first-semester, five-credit calculus course in the spring of 1973 could elect self-paced instruction. One-hundred-five students chose the self-paced section. These 105 students are compared here with about 130 students who were in a traditionally run section. Different instructors were responsible for the two groups, but there was a single, common final examination.

Those electing traditional instruction attended three 50–minute lectures and two 50–minute recitations weekly. Quizzes and examinations were administered, graded and returned in the traditional manner.

Students electing self-paced instruction could select any of three differently-paced presentations — fast, medium, and slow; after a few weeks, these were modified to one medium and two slow sections. Classes met in four 50–minute sessions per week, and students could choose to attend any of the differently-paced sections offered on a given day. In addition, one and one-half days per week were provided for taking the 12 quizzes and the midterm examination. Both groups used the same text: Leithold, *The Calculus with Analytic Geometry* [2]. The syllabus included most of the first six chapters.

A student had the option of taking each quiz up to five times, and only the best score contributed to the student's grade. Each quiz was graded within a day. The student then discussed any errors or misunderstandings with an instructor and received individual help with any difficulties.

Students from the two groups took the same final examination simultaneously. Some students in the self-paced group were given the option of taking an earlier final and exempting the common final. Fifteen took this early final; eleven of them were exempted from the common final. The comparison of the two groups was made on the basis of scores on the common final examination. The criteria for the grading of the final were standardized in joint discussions among the staff members responsible for the two groups, but each group was graded by its own staff members.

One author (R. R. S.) designed and was the primary instructor in the self-paced section; the other (R. J. F.) designed and conducted the statistical analysis.

Results — Table I presents the distribution of final examination scores. Recall that scores of eleven students (19%) from the self-paced section are not included in this distribution because of their exemption from the final examination.

Of interest in Table II, Means of Final Examination Scores, is a comparison of the mean of scores of the self-paced section with the mean of scores of the traditional section. This difference of 172.7 *vs.* 153.3 is significant at the 0.01 level, ($F_{1,148} = 8.8$; both grade point average and prior mathematical ability taken into account; details of all statistical analyses are given in the appendix). While the

TABLE I. *Distribution of Final Examination Scores*

Final Exam Score*	Letter Grade	Self-paced No.	Self-paced %	Self-paced Cum. %	Traditional No.	Traditional %	Traditional Cum. %
190–200	A	11	19.2	19.2	18	18.6	18.6
180–189		14	24.6	43.8	16	16.5	35.1
170–179	B	12	21.1	64.9	9	9.3	44.4
160–169		6	10.5	75.4	10	10.3	54.7
150–159	C	6	10.5	85.9	6	6.2	60.9
140–149		5	8.8	94.7	10	10.3	71.2
130–139	D	1	1.8	96.5	7	7.2	78.4
120–129		0	0	96.5	5	5.2	83.6
110–119	F	2	3.5	100.0	4	4.1	87.7
100–109		0	0	100.0	0	0	87.7
Below 100		0	0	100.0	12	12.4	100.1
Total		57	100.0		97	100.1	

* Maximum Score = 200

difference in scores of the males (157.9) and females (167.5) is quite large, it is not significant ($\alpha = 0.05$). The data summarized in Table II were also analyzed for a sex by method interaction; i.e., for evidence that one method or the other was more effective for one sex. The interaction was not significant ($\alpha = 0.05$). Final examination scores were also analyzed for a method by ability interaction to see if one method was better for students of a given ability level. Again, the interaction was not significant ($\alpha = 0.05$). From these analyses we conclude ($p = 0.99$) that the differences in the examination scores of the self-paced and traditional groups can be attributed to method. We find no reason to believe that self-paced instruction (or traditional instruction) is better for one sex than for the other, or better for students of one ability level than the others.

In Table III, Means of Cumulative Grade Point Averages, a comparison between the grade point averages (GPA) of the self-paced group (2.90) and the traditional group (2.77) shows that the averages are not significantly different. The difference between the GPA of the males (2.72) and the females (3.10) is significant at the 0.01 level ($F_{1,150} = 9.8$).

Table IV gives means of scores on a test of prior mathematical ability, the American Testing Program's mathematics examination. (ATP is comparable to the Educational Testing Service's Scholastic Aptitude Test (SAT) in mathematics.) These scores favor the traditional group (60.4) over the self-paced group (57.9). This difference is significant at the 0.10 level ($F_{1,150} = 3.1$). This difference was compensated for by using ATP score as a covariate when analyzing the final examination scores. The difference between males (59.8) and females (58.6) is not statistically significant ($\alpha = 0.05$).

Since no student in the self-paced group failed, it might be argued that the failing students in the traditional group might spuriously have depressed the performance of that group. These students might have been counselled to withdraw from the course had they received the regular counselling

TABLE II. *Means of Final Examination Scores**

		Sex		Row Means
		Male	Female	
Method	Self-paced	$\bar{X} = 172.9$ $n = 41$	$\bar{X} = 172.3$ $n = 16$	$\bar{X} = 172.7$ $n = 57$
	Traditional	$\bar{X} = 149.4$ $n = 72$	$\bar{X} = 164.5$ $n = 25$	$\bar{X} = 153.3$ $n = 97$
	Column Means	$\bar{X} = 157.9$ $n = 113$	$\bar{X} = 167.5$ $n = 41$	

* Maximum Score = 200

TABLE III. *Means of Cumulative Grade Point Averages**

		Sex		Row Means
		Male	Female	
Method	Self-paced	$\bar{X} = 2.83$ $n = 41$	$\bar{X} = 3.07$ $n = 16$	$\bar{X} = 2.90$ $n = 57$
	Traditional	$\bar{X} = 2.65$ $n = 72$	$\bar{X} = 3.12$ $n = 25$	$\bar{X} = 2.77$ $n = 97$
	Column Means	$\bar{X} = 2.72$ $n = 113$	$\bar{X} = 3.10$ $n = 41$	

* A = 4.0, B = 3.0, C = 2.0, D = 1.0

TABLE IV. *Means of ATP* Mathematics Examination Scores*

		Sex		Row Means
		Male	Female	
Method	Self-paced	$\bar{X} = 58.1$ $n = 41$	$\bar{X} = 57.3$ $n = 16$	$\bar{X} = 57.9$ $n = 57$
	Traditional	$\bar{X} = 60.8$ $n = 72$	$\bar{X} = 59.4$ $n = 25$	$\bar{X} = 60.4$ $n = 97$
	Column Means	$\bar{X} = 59.8$ $n = 113$	$\bar{X} = 58.6$ $n = 41$	

* American Testing Program, Mean = 50, SD = 10, Maximum Score = 80

available to students in the self-paced group. The data were, therefore, reanalyzed, deleting from the sample the scores of students in the traditional group with a final grade of F. Mean scores for this analysis are given in Table V.

Even with the scores of failing students in the traditional group deleted, the performance of the self-paced group remains significantly higher ($F_{1,135} = 5.7$), although at a slightly higher α of 0.025. The conclusion remains ($p = 0.975$) that differences in examination scores were caused by the different methods of instruction. There is no significant difference ($\alpha = 0.05$) between scores of males and females.

TABLE V. *Cell Means for Final Examination* with Scores of Failing Students Deleted*

		Sex		Row Means
		Male	Female	
Method	Self-paced	$\bar{X} = 172.9$ $n = 41$	$\bar{X} = 172.3$ $n = 16$	$\bar{X} = 172.7$ $n = 57$
	Traditional	$\bar{X} = 164.9$ $n = 60$	$\bar{X} = 168.3$ $n = 24$	$\bar{X} = 165.9$ $n = 84$
	Column Means	$\bar{X} = 168.1$ $n = 101$	$\bar{X} = 169.9$ $n = 40$	

* Maximum Score = 200

TABLE VI. *Students who did not pass the Course*

Reason	Self-paced		Traditional	
	No.	%*	No.	%**
Grade of F***	0	0	13	9.8
Withdrew	21	20.0	16	12.1
Grade of Incomplete	16	15.2	19	14.4
Total	37	35.2	48	36.3

* Based on 105 who started
** Based on 133 who started
*** Deleted in Table V

During the semester, forty-five students in the self-paced group replied to a questionnaire concerning the course. Of these 64% preferred this method over a traditional mathematics class, and 75% thought that the Department of Mathematics should offer a few self-paced sections in the following year. Some of their positive comments were: learned a lot; liked the personal attention and contact; liked finishing early; liked having several attempts at each quiz; liked having a choice of instructor. Some of their complaints: too many tests; grading too stringent; too time-consuming; too easy to fall behind.

Discussion — The conclusion to be drawn from this analysis is quite clear: students in the self-paced group did significantly better ($p = 0.975 - 0.99$) than students receiving traditional instruction. This is so even though some top students in the self-paced group (those who exempted the common final) are omitted from the analysis, a factor that favors the traditional group. Even when scores of failing students (all of whom were in the traditional group) are deleted from the analysis, the scores of the self-paced group remain significantly higher at the 0.025 level. These results indicate strongly the superiority of the self-paced method.

About a third of the students in each group did not pass the course for one of several reasons. These students either withdrew, received grades of incomplete, or failed the course. Table VI gives the distribution of students in each of these categories.

It is important to analyze these withdrawal rates because it is possible, inadvertently, to design a self-paced course which eliminates all but the top students. These students would then obviously outperform the mixture of students that would remain in a traditional class. There were, indeed, more withdrawals and incompletes among the self-paced group; but there were no failures. The percentages of students who, finally, did pass are, however, about the same in the two groups. One might argue

that students who were doing poorly in the self-paced course had more, and more regular, contact with instructors who might have encouraged them to withdraw or to take an incomplete, while in the traditional group, some of these students might not have been identified until the very end of the semester. There is no evidence that the self-paced method eliminated more students than the traditional.

It is recognized that there are several flaws in the design of this experiment, all of which reflect the fact that the analysis was undertaken only after the course had been given. (1) Students were not randomly assigned to the two groups but had to select the self-paced group. (Although the design of the experiment might have been improved by random assignment, for pedagogical reasons, the authors believe that self-pacing should be available only to those who elect it.) (2) Although the criteria for grading the common final examination were agreed on in advance by the instructors in the course, each staff graded only the papers of its own group. (3) Each instructor taught by only one method. (One instructor assisted by two teaching assistants taught the self-paced group, and a different instructor assisted by three different teaching assistants taught the traditional group.) The possibility that the observed differences might be the result of the different instructors rather than the different methods cannot be excluded. There are, however, two reasons for believing that it is the method rather than the instructors that underlies the differences. First, the instructor of the traditional class is generally recognized as a fine teacher.

Second, the subjective feeling of the instructor of the self-paced section (R. R. S.) is that the students learned more than students taught by her in calculus classes conducted in a traditional manner. A very likely explanation for the difference is the greater time spent by the instructor with each student, a factor that is intrinsic to the self-paced method. Some students received what amounted to extensive tutoring by some of the instructors in the self-paced section. It may be that any method that utilizes so large an amount of instructional time will give similar results.

Two implications for departments of mathematics can be drawn from this experiment. First, since the students who chose the self-paced section achieved better scores than students in the traditional section, this type of instruction should, whenever possible, be made available to students who desire it. Departments of mathematics should also be aware of the greater instructional time involved in self-paced instruction and should be prepared to assign additional faculty to make this option available.

Second, data derived during this study argue strongly that the women who attempted this beginning course in the calculus may have been even more competent than their male peers: the sexes were not different with respect to ATP scores; the females appeared to outperform the males on the common final examination; and the grade point averages of females in the course clearly exceeded those of the males. In a period of declining enrollments in upper division courses, departments of mathematics should recognize females as a potentially rich source of capable mathematics majors.

Appendix — The final examination scores were analyzed by means of a two-way analysis of covariance (ANCOVA) model in which the factors method and sex were crossed. Computer program BMD 05V [1] was employed for the ANCOVA, covarying on mathematical ability (as measured by ATP score in mathematics) and college cumulative average (GPA) at the end of the semester. In a second analysis, using the same design, scores of failing students were deleted. Again, ATP scores and GPA were covariates.

A second model, primarily for testing ability × method interaction, was also employed. The model was a three-way analysis of variance (ANOVA), crossing the factors of mathematical ability, method, and sex. Computer program BMD 05V was again used, but without covariates.

Cumulative grade point average and ATP scores in mathematics were analyzed in a two-way ANOVA in which the factors method and sex were crossed.

The authors wish to thank Dr. Abraham S. Flexer for his constructive criticism of this paper.

References

1. W. J. Dixon, BMD Biomedical Computer Programs, University of California Press, Berkeley, 1973.
2. L. L. Leithold, The Calculus with Analytic Geometry, Harper & Row, New York, 1969.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COLORADO, BOULDER, CO80302.
SCHOOL OF EDUCATION, UNIVERSITY OF COLORADO, BOULDER, CO80302.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before May 31, 1977.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

E 2635. *Proposed by Kirby C. Smith, Texas A & M University*

Let F be a field of characteristic $p \neq 0$. Let $A = CD$ where C is a cyclic $p \times p$ matrix over F and D is the diagonal matrix with diagonal entries $0, 1, 2, \ldots, p-1$. Compute the characteristic polynomial of A. Generalize.

E 2636. *Proposed by D. E. Knuth, Stanford University*

A pair of microbes was recently discovered which reproduce in a very peculiar way. The male microbe (a diphage) has two receptors on its surface, and the female (a triphage) has three receptors. When a culture of diphages and triphages is irradiated with a psiparticle, exactly one of the receptors absorbs the particle (each receptor being equally likely). If it was a diphage, it changes to a triphage; but if it was a triphage, it splits into two diphages.

Give a simple formula for the average number of diphages present if we begin with a single diphage and irradiate the culture n times with psi-particles.

E 2637. *Proposed by Armond E. Spencer, State University College, Potsdam, N.Y.*

If $a_0, a_1, \ldots, a_{n-1}$ are integers show that

$$\prod_{0 \leqq i < j \leqq n-1} \frac{a_i - a_j}{i - j}$$

is also an integer.

E 2638. *Proposed by Robert McNaughton, Rensselaer Polytechnic Institute*

Call a set of positive integers a *clique* if no two of its elements are relatively prime. Call a member of a clique a *leader* if it is not a proper multiple of another member of the clique. Construct a maximal clique with infinitely many leaders. (The set of all cliques is partially ordered by inclusion.)

E 2639. *Proposed by G. Tsintsifas, Thessaloniki, Greece*

Let ABC be a triangle with $\measuredangle A = 40°$, $\measuredangle B = 60°$. Let D and E be points lying on the sides AC and AB, respectively, such that $\measuredangle CBD = 40°$ and $\measuredangle BCE = 70°$. Let F be the point where the lines BD and CE intersect. Show that the line AF is perpendicular to the line BC.

E 2640*. *Proposed by James E. Desmond and William R. Hastings, Pensacola Junior College, Florida*

Prove or disprove: The largest power of 2 which divides

$$\binom{2^{n+1}}{2^n} - \binom{2^n}{2^{n-1}}, \qquad (n > 1)$$

is 2^{3n}.

SOLUTIONS OF ELEMENTARY PROBLEMS

Indefinite Quadratic Form on a Box

E 2555 [1975, 851]. *Proposed by T. W. Cusick, University of Illinois*

Let $A = (\mathbf{a}_1 \,|\, \mathbf{a}_2)$ be a nonsingular 2×2 matrix partitioned into columns. Show that

$$\min_A \max_x \frac{(\mathbf{a}_1 \cdot \mathbf{x})(\mathbf{a}_2 \cdot \mathbf{x})}{\det A} = \tfrac{1}{2},$$

where the max is over all $\mathbf{x}$ in the box $|x_i| \leqq 1$, and the min is over all such matrices A.
(*) Establish a corresponding result for higher dimensions.

I. *Solution* (*to the problem as printed*) *by M. J. Pelling, University of Benin, Nigeria.* The question is wrong as stated since with $A = \begin{pmatrix} 1 & -\alpha \\ 0 & 1 \end{pmatrix}$, $\alpha > 0$, the max is $1/4\alpha$ on the box, which can be made as close to zero as we please.

II. *Solution* (*to the problem as stated originally by the proposer*) *by Vinod Kumar Grover, Panjab University, Chandigarh, India.* Let $f(A, \mathbf{x}) = (\mathbf{a}_1 \cdot \mathbf{x})\ (\mathbf{a}_2 \cdot \mathbf{x})/\det A$ and let Δ be the box $|x_i| \leqq 1$. We prove that

$$\min_A \max_{\mathbf{x} \in \Delta} |f(A, \mathbf{x})| = \tfrac{1}{2}.$$

If $A_0 = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ then $\max_{\mathbf{x} \in \Delta} |f(A_0, \mathbf{x})| = 1/2$ and it suffices to show that for every non-singular matrix A we have

$$\max_{\mathbf{x} \in \Delta} |f(A, \mathbf{x})| \geqq 1/2.$$

Equivalently, it suffices to show that if A is a non-singular matrix and

$$(1) \qquad \max_{\mathbf{x}\in\Delta} |(\mathbf{a}_1\cdot\mathbf{x})(\mathbf{a}_2\cdot\mathbf{x})| = 1/2,$$

then $|\det A|\leqq 1$. Suppose that (1) holds and that $|\det A|>1$. The set of points $(\mathbf{a}_1\cdot\mathbf{x},\mathbf{a}_2\cdot\mathbf{x})$ for $\mathbf{x}\in\Delta$ is a parallelogram Π with center O and area $4|\det A|>4$. By (1) Π is contained in the region $|uv|\leqq 1/2$ where u, v are the coordinates in $\mathbf{R}^2$. Let $P=(0,\sigma)$ with $\sigma>0$ be the boundary point of Π. Let Λ be the lattice generated by the vectors $(1/\sigma,\sigma/2)$ and $(-1/\sigma,\sigma/2)$. By Minkowski's theorem on convex bodies Π must contain a non-zero point Q of Λ in its interior, say

$$Q = m(1/\sigma,\sigma/2)+n(-1/\sigma,\sigma/2) = ((m-n)/\sigma,(m+n)\sigma/2),$$

where $m, n\in\mathbf{Z}$ and we may assume that $(m,n)=1$. Since the interior of Π lies in the region $|uv|<1/2$ we must have $|m^2-n^2|<1$. Therefore $m^2=n^2=1$, i.e., the possible points Q are $(0,\pm\sigma)$, $(\pm 2/\sigma,0)$. But $(0,\pm\sigma)$ are on the boundary of Π and consequently we can take $Q=(2/\sigma,0)$. The midpoint $S=(1/\sigma,\sigma/2)$ of PQ lies also in the interior of Π. This is a contradiction because S lies on the curve $|uv|=1/2$.

Also solved (first version) by Gary Bates, Albert Briggs, Jr., Paul Bruckman, G. A. Heuer & Karl Heuer, and Paul Vojta.

Solutions to the second version were submitted also by M. J. Pelling (Nigeria), and the proposer.

Editor's Comment. No solution for higher dimensions was received. Grover claims (without proof) that for matrices of order n the constant 1/2 should be replaced by $2^n/V$ where V is the volume of the largest parallelepiped with center O contained in the region $|u_1u_2\cdots u_n|\leqq 1$ of $\mathbf{R}^n$. Solutions of the generalization (in the second, corrected, version of the problem) are still solicited.

A Rank Argument

E 2556 [1975, 852]. *Proposed by Leon Gerber, St. John's University*

Let $A=(\mathbf{a}_1|\cdots|\mathbf{a}_n)$ and $B=(\mathbf{b}_1|\cdots|\mathbf{b}_n)$ be two $2n\times n$ real matrices, partitioned into columns. Assume that $n\geqq 3$ and that the rank of A does not exceed $n-3$. Let $r_1,\ldots,r_n, s_1,\ldots,s_n$ be arbitrary positive numbers. For $i,j=1,2,\ldots,n$, define

$$t_{ij}=\frac{|\mathbf{a}_i-\mathbf{b}_j|^2-r_i^2-s_j^2}{2r_is_j}.$$

Show that $\det(t_{ij})=0$.

Solution by A. A. Jagers, Technische Hogeschool Twente, Enschede, Netherlands. Let $\alpha = 2^n r_1\cdots r_n s_1\cdots s_n$. Then

$$\alpha\det(t_{ij})=\det(X-2A'B),$$

where $X=(x_{ij})$ and $x_{ij}=|\mathbf{a}_i|^2+|\mathbf{b}_j|^2-r_i^2-s_j^2$. The difference of any two rows of X is a scalar multiple of $(1,\ldots,1)$ and hence $\operatorname{rank}(X)\leqq 2$. Since $\operatorname{rank}(2A'B)\leqq\operatorname{rank}A\leqq n-3$ it follows that $\operatorname{rank}(X-2A'B)\leqq n-1$ and consequently $\det(t_{ij})=0$.

Also solved by Christopher Henley, L. S. de Jong (Netherlands), Harry Lass, and the proposer.

"Perfect" Cyclic Quadrilaterals

E 2557 [1975, 852]. *Proposed by R. D. Nelson, Ampleforth College, York, England*

Find all cyclic quadrilaterals with integral sides, each of which has its perimeter numerically equal to its area. (The following references may be of interest: E 1168 [1956, 43]; E 2420 [1974, 662]; M. V. Subbarao, *Perfect triangles*, this MONTHLY 78 (1971), 384–385; R. W. Sielaff, *Perfect quadrilaterals*, this

	x	y	z	t	a	b	c	d	R
(i)	1	9	10	10	14	6	5	5	$\frac{5}{6}\sqrt{109}$
(ii)	2	5	5	8	8	5	5	2	$\frac{5}{8}\sqrt{41}$
(iii)	3	3	6	6	6	6	3	3	$\frac{3}{2}\sqrt{5}$
(iv)	4	4	4	4	4	4	4	4	$2\sqrt{2}$

Each of the cases (i), (ii), (iii) gives two cyclic quadrilaterals by choosing two different ways of arranging the sides. Hence there are seven quadrilaterals which have all the required properties.

The last column of the above table gives the radius R of the circle on which the four vertices of the quadrilateral lie.

Also solved by Temple University Problem Solving Group and by the proposer.

A Theorem of Dini

E 2558 [1975, 936]. *Proposed by A. Torchinsky, Cornell University*

Suppose that Σa_n is a divergent series of positive terms, and let $s_n = a_1 + \cdots + a_n$ for $n = 1, 2, \ldots$. For which values of p does the series $\Sigma a_n/s_n^p$ converge? (Remark: The special cases $p = 1$ and $p = 2$ are problem 11, p. 70, of Walter Rudin, *Principles of Mathematical Analysis* (Second Edition), McGraw–Hill, New York, 1964.)

I *Solution by Leon Gerber, St. John's University, Jamaica, N.Y.* If $p > 1$, then

$$\frac{a_k}{s_k^p} = \frac{s_k - s_{k-1}}{s_k^p} < \int_{s_{k-1}}^{s_k} \frac{dx}{x^p}$$

and the given series converges since $\int_1^\infty dx/x^p$ converges.

If $p \leqq 1$, then for large k we have $s_k > 1$ and

$$\frac{a_k}{s_k^p} = \frac{s_k - s_{k-1}}{s_k^p} \geqq \frac{s_k - s_{k-1}}{s_k} = 1 - \frac{s_{k-1}}{s_k}.$$

Hence the given series diverges since the product $\Pi_{k=2}^\infty(s_{k-1}/s_k)$ diverges to zero.

II *Generalization by M. J. Pelling, University of Benin, Nigeria.* We shall prove the following two theorems.

THEOREM 1. *Let $f(x) > 0$ for $x > 0$. Then $\Sigma a_n/f(s_n) = \infty$ for all positive divergent series Σa_n if and only if $f(x) = O(x)$ as $x \to \infty$.*

THEOREM 2. *Let $g(x) \geqq 0$ for $x \geqq 0$ and let $h(x) = \sup_{y \geqq x} g(y) \leqq \infty$. Then $\Sigma a_n g(s_n) < \infty$ for all positive divergent series Σa_n if and only if there exists $\alpha \geqq 0$ such that $\int_\alpha^\infty h(x)dx < \infty$.*

Proof of Theorem 1. Let $f(x) < Ax$ for $x > x_0$ and $s_k > x_0$. Then

$$\sum a_n/f(s_n) \geqq A^{-1} \sum_{n \geqq k} a_n/s_n = \infty.$$

If $f(x) \neq O(x)$ $(x \to \infty)$ then $\liminf x/f(x) = 0$ $(x \to \infty)$. Hence we can define $s_1 < s_2 < \cdots$ $(s_n \to \infty)$ so that

$$\frac{a_n}{f(s_n)} = \frac{s_n - s_{n-1}}{s_n} \cdot \frac{s_n}{f(s_n)} < 2^{-n}.$$

Then $\Sigma\, a_n = \infty$ but $\Sigma\, a_n/f(s_n) < \infty$.

Proof of Theorem 2. Suppose $\int_\alpha^\infty h(x)dx < \infty$ and let $s_{k-1} > \alpha$. Then

$$\sum_{n\geqq k} a_n g(s_n) \leqq \sum_{n\geqq k} a_n h(s_n) = \sum_{n\geqq k} (s_n - s_{n-1})h(s_n) \leqq \int_{s_{k-1}}^\infty h(x)dx < \infty,$$

so that $\Sigma\, a_n g(s_n)$ converges.

Conversely, suppose that $\Sigma\, a_n g(s_n) < \infty$ for all positive divergent series $\Sigma\, a_n$. If $\limsup xg(x) = A > 0$ $(x \to \infty)$ we can define $s_1 < s_2 < \cdots$ $(s_n \to \infty)$ such that

$$a_n g(s_n) = \frac{s_n - s_{n-1}}{s_n} \cdot s_n g(s_n) > \tfrac{1}{2}A.$$

This would give $\Sigma\, a_n g(s_n) = \infty$ contrary to hypothesis. Therefore $xg(x) \to 0$ $(x \to \infty)$ and g and h are bounded near ∞. Since the convergence of $\Sigma\, a_n g(s_n)$ is not affected by changing g in a finite interval we can assume that g has been altered so that g and h are bounded in $[0, \infty)$.

Assuming that $\int_0^\infty h(x)dx = \infty$ there is a partition $0 = t_0 < t_1 < \cdots$ $(t_n \to \infty)$ of $[0, \infty)$ such that $\Sigma(t_n - t_{n-1})h(t_n) = \infty$. We may further assume that $h(t_n) < h(t_{n-1})$ for $n \geqq 1$ (by lumping together several consecutive intervals, if necessary).

We can choose s_n such that $t_n \leqq s_n < t_{n+1}$ and

$$g(s_n) > \max\left(\tfrac{1}{2}h(t_n), h(t_{n+1})\right).$$

Put $s_0 = 0$ and $a_n = s_n - s_{n-1}$ $(n \geqq 1)$ so that $\Sigma\, a_n$ is divergent and

$$\begin{aligned}\sum a_n g(s_n) &= \sum (s_n - t_n)g(s_n) + \sum (t_n - s_{n-1})g(s_n)\\ &\geqq \tfrac{1}{2}\sum (s_n - t_n)h(t_{n+1}) + \tfrac{1}{2}\sum (t_n - s_{n-1})h(t_n)\\ &= \tfrac{1}{2}\sum (t_n - t_{n-1})h(t_n) = \infty.\end{aligned}$$

This contradicts our hypothesis, so, after all, $\int_0^\infty h(x)dx < \infty$.

Solutions similar to I were submitted by Raymond Chen, Michael Dixon & Hugh Noland, and Richard Groeneveld & Dean Isaacson. Also solved by S. N. Ethier, William Habakkuk, Elmer Hayashi, Ellen Hertz, Simeon Reich, Samson Rosenzweig, Tavan Trent, Barbara Turner, Temple University Problem Group, University of South Alabama Problem Group, David Wright, and the proposer. A partial solution was submitted by Irving Dodes.

The following contributors point out that the result can be found in the literature, for example in K. Knopp, *Infinite Sequences and Series*, Dover, New York, 1956, p. 125, where it is attributed to Dini: R. C. Carson, Paul Chauveheid (Belgium), Charles Chouteau, G. P. Farrell, Max Garbutt & Roger Walters, P. K. Garlick, M. R. Gopal, Michael Hoffman, Richard Johnsonbaugh, Murray Klamkin (Canada), Wm. R. Klinger, J. C. Lagarias, O. P. Lossers (Netherlands), Robert McCarthy, A. Meir (Canada), H. Niederreiter, Andris Niedra, T. Šalát (Czechoslovakia), Bruce Shawyer (Canada), and W. C. Waterhouse.

A Nonsingular Matrix

E 2559 [1975, 936]. *Proposed by Hugh L. Montgomery, University of Michigan*

Determine whether the following matrix is singular or nonsingular:

$$\begin{pmatrix} 51237 & 79922 & 55538 & 39177\\ 46152 & 16596 & 37189 & 82561\\ 71489 & 23165 & 26563 & 61372\\ 44350 & 42391 & 91185 & 64809 \end{pmatrix}.$$

Solution by Ralph Howard and Lorraine L. Foster, California State University at Northridge. The matrix is congruent mod 5 to the matrix

$$\begin{pmatrix} 2 & 2 & -2 & 2 \\ 2 & 1 & -1 & 1 \\ -1 & 0 & -2 & 2 \\ 0 & 1 & 0 & -1 \end{pmatrix}$$

which has determinant 1 (mod 5). Hence the original matrix is nonsingular.

Also solved by the proposer and 68 other readers.

Non-congruence of Certain Sums

E 2560 [1975, 936]. *Proposed by Richard Madsen, University of Missouri at Columbia*

Let $n_1, \ldots, n_k$ be natural numbers. Define $d_1 = 1$ and $d_i = \mathrm{GCD}(n_1, \ldots, n_{i-1})/\mathrm{GCD}(n_1, \ldots, n_i)$ for $i \geqq 2$. Show that the $d_1 \cdots d_k$ possible sums

$$\sum_{i=1}^{k} a_i n_i \qquad a_i \in \{1, 2, \ldots, d_i\}$$

are all distinct modulo n_1.

Solution by University of South Alabama Problem Group. Suppose that for two of these sums we have

$$\sum_{i=1}^{k} a_i n_i - \sum_{i=1}^{k} b_i n_i = n_1 u \qquad (u \in \mathbb{Z}). \tag{1}$$

Let s be the largest integer such that $a_s \neq b_s$ and let $n_s = c \cdot \mathrm{GCD}(n_1, \ldots, n_s)$. Dividing (1) by $\mathrm{GCD}(n_1, \ldots, n_s)$ we see that d_s divides $(a_s - b_s)c$. Since $\mathrm{GCD}(d_s, c) = 1$ we obtain $d_s \mid (a_s - b_s)$, which is a contradiction because $1 \leqq |a_s - b_s| < d_s$.

Also solved by Paul Bruckman, Martin Flashman, Lorraine Foster, Donald Fuller, M. G. Greening (Australia), Frederick Humburg, Carl Hurd, Bhushan Lal Kapoor (India), L. Kuipers (Switzerland), Barry Light, Louise Moser, St. Olaf Problem Group, Robert Patenaude, Temple University Problem Solving Group, A. Tietäväinen (Finland), and the proposer.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Hill Center, Rutgers University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate signed sheets and should be mailed before May 31, 1977.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

6132. *Proposed by Mihai Eşanu, Liceul "Gheorghe Lazăr," Bucharest, Romania*

Find all the functions $f: R \to R$ with the Darboux property such that for some $n \geqq 1$, $f^n(x) = -x$ for all x. (Here $f^2 = f \circ f$, etc.)

6133. *Proposed by J. Cano and A. Gruebler, Universidad Simon Bolivar, Venezuela*

Let $f: [a, b] \to [a, b]$ be continuous and denote $P(f) = \{x : f^n(x) = x$ for some $n = 1, 2, \ldots\}$, $C(f) = \{x : f^m(x) \in P(f)$ for some $m = 1, 2, \ldots\}$, and L_x the set of limit points of the sequence

$\overline{\{x, f(x), f^2(x), \ldots, f^n(x), \ldots\}}$. Here $f^n(x)$ is the nth iterate of $f(x)$. Prove that for each $x \in [a, b]$, $L_x \subset C(f)$. Is this true in R^2?

6134. *Proposed by Barbara Osofsky, Rutgers University*

Let R be a ring, not necessarily with identity, and let R^n be the subring generated by n-fold products of elements of R. Prove: If R has descending chain condition (d.c.c.) on right ideals, then so does R^n. Does this result hold if d.c.c. is replaced by ascending chain condition (a.c.c.)?

6135*. *Proposed by Paul Erdös, Hungary*

Denote by $P(n)$ the greatest prime factor of n and put $A(x, y) = \Pi_{1 \leqq i \leqq y-x}(x + i)$. An integer n is called *exceptional* if for some $x \leqq n \leqq y$, $(P(A(x, y)))^2$ divides $A(x, y)$, i.e., the greatest prime factor of $A(x, y)$ occurs with an exponent greater than one.

Prove that the density of exceptional numbers is 0 and estimate the number $E(x)$ not exceeding x as well as you can.

6136. *Proposed by H. L. Montgomery, University of Michigan*

Let $P(z, w) = \Sigma c_{mn} z^m w^n$ be a polynomial in $\mathbb{C}[z, w]$. Suppose that $Q(z, w) = P(z, w/z)$ is also a polynomial: that is $c_{mn} = 0$ whenever $n > m$. Show that

$$\{P(z, w): |z| < 1, |w| < 1\} = \{Q(z, w): |z| < 1, |w| < 1\}.$$

6137*. *Proposed by I. J. Good, Virginia Polytechnic Institute and State University*

Let $p(n)$ denote the number of partitions of n ($n = 1, 2, \ldots$), and let k denote an integer greater than 3. Prove that $\Delta^k p(n)$ ($n = 1, 2, \ldots$) is a sequence of alternating terms.

SOLUTIONS OF ADVANCED PROBLEMS

Distance to the Boundary of a Set

6025 [1975, 409]. *Proposed by S. F. Wong and B. B. Winter, University of Ottawa*

Let (X, d) be a metric space, T an arbitrary subset of X, and t an arbitrary element of T. As usual, $d(t, A) = \inf\{d(t, a): a \in A\}$ is $-\infty$ if $A = \emptyset$; ∂T and T^c are, respectively, the boundary and the complement of T.

(a) Is it always true that $d(t, x) < d(t, \partial T)$ implies $x \in T$?

If not, find a condition on (X, d) which is necessary and sufficient for the validity of this implication.

(b) Is it always true that $d(t, \partial T) = d(t, T^c)$?

If not, find a condition of (X, d) which is necessary and sufficient for the validity of this equality.

Solution by C. Bruce Hughes, Guilford College. Conditions (a) and (b) are equivalent, and are also equivalent to:

(c) Every ε-neighborhood $N_\varepsilon(x)$ is connected.

Proof. (a) implies (b). Using the triangle inequality it is easy to verify that $d(t, T^c) \leqq d(t, \partial T)$ is true in any metric space. If $d(t, T^c) < d(t, \partial T)$, then there exists $x \in T^c$ with $d(t, x) < d(t, \partial T)$. It follows from (a) that $x \in T$, which contradicts $x \in T^c$.

(b) implies (c). Suppose that there exists an ε-neighborhood $N_\varepsilon(x)$ which is not connected. Let $N_\varepsilon(x) = A \cup B$ be a separation, and assume that $x \in A$. Since A and B are open, $\partial A \cap N_\varepsilon(x) = \partial A \cap (A \cup B) = \emptyset$. By (b) $\partial A \neq \emptyset$, hence $d(x, \partial A) = \varepsilon$. Since $B \subseteq A^c$, it follows that $d(x, A^c) \leqq d(x, B) < \varepsilon$. But $d(x, \partial A) = d(x, A^c) < \varepsilon$.

(c) implies (a). Suppose that there exists $x \in T^c$ and $t \in T$ such that $d(t, x) < d(t, \partial T)$. Then there

exists $\varepsilon > 0$ such that $x \in N_\varepsilon(t)$ and $N_\varepsilon(t) \cap \partial T = \emptyset$. Since $N_\varepsilon(t)$ is connected and intersects both T and T^c, $N_\varepsilon(t) \cap \partial T \neq \emptyset$.

Also solved by Bethany College Problems Group, David Browder, Victor Manjarrez & Louise Moser, David Ritter, Thomas Sellke, Carlton Woods, and the proposers.

Editor's Note. Several solvers pointed out that $d(x, \emptyset)$ is usually taken to be $+\infty$ rather than $-\infty$. This convention is required in the above solution.

Number of Elements in a Group Inverted by an Automorphism

6026 [1975, 409]. *Proposed by Fred Commoner, Cambridge, Massachusetts*

Prove the theorem: Let p be an odd prime. If G is a finite non-abelian group such that p is less than or equal to the least prime dividing $|G|$, then no automorphism of G can send more than $|G|/p$ elements of G to their inverses. There is a non-abelian group G of order p^3 and an automorphism of G sending exactly $|G|/p$ elements of G to their inverses.

Solution by A. A. Jagers, Technische Hogeschool Twente, Enschede, Netherlands. (1) If $|G|$ is even, then there is nothing to prove, so suppose $|G|$ is odd. Let $\sigma \in \mathrm{Aut}(G)$ and put $S = \{x \in G \mid \sigma x = x^{-1}\}$, $C_i = \{x \in G \mid \sigma^i x = x\}$, $i = 1, 2$. Then C_2 contains S and $|C_2|$ divides $|G|$ by Lagrange's theorem. Now suppose by way of contradiction that $|S| > |G|/p$. Then $C_2 = G$; i.e., σ has order two. Since in addition G is non-abelian, σ cannot be fixed-point-free (see [**1**], p. 336, thm. 1.4). That is $|C_1| > 1$, and by our assumption on $|S|$ we have $|S| > [G: C_1]$. Hence at least one coset of C_1 contains more than one element of S. Say, $x, y \in S$ with $x = yc$, $c \in C_1$, $c \neq e$. Then $c^{-1}y^{-1} = x^{-1} = \sigma x = (\sigma y)(\sigma c) = y^{-1}c$; or $y^{-1}cy = c^{-1}$. Consequently $y^{-2}cy^2 = c$. Hence, either $c^{-1} = c$, or the centralizer of c in $\langle y \rangle$ is a proper subgroup of $\langle y \rangle$ generated by y^2. However, $|G|$ is odd, so the order of both c and y is odd: a contradiction in either case.

(2) Up to isomorphism there are precisely two nonabelian groups of order p^3 (cf. [**2**], p. 93), viz.

$$G = \langle n, s \mid s^{-1}ns = n^{1+p}, n^{p^2} = s^p = e \rangle,$$

resp.

$$G = \langle a, b, c \mid a^p = b^p = c^p = e,\ a^{-1}ba = bc,\ a^{-1}ca = b^{-1}cb = c \rangle.$$

In either case G is a semidirect product of an abelian normal subgroup N and a p-cyclic subgroup, and so an automorphism ϕ with the desired properties can be obtained by extending the isomorphism $x \to x^{-1}$ $(x \in N)$ to G. Explicitly, put $\phi n = n^{-1}$, $\phi s = s$ in the first case and $\phi a = a$, $\phi b = b^{-1}$, $\phi c = c^{-1}$ in the second case.

References

1. D. Gorenstein, Finite Groups, Harper & Row, New York, 1968.
2. B. Huppert, Endliche Gruppen I, Springer, Berlin, 1967.

Also solved by Peter Borwein & Martin Schechter (Canada), Allan Coppola, M. G. Greening (Australia), Kenneth Klinger, O. P. Lossers (Netherlands), R. C. Lyndon, Victor Manjarrez & Louis Moser, M. R. Modak (India), Ram Murty & Kumar Murty (Canada), Barbara Osofsky, S. G. Updikar (India), and the proposer.

Editor's Comment. Desmond MacHale (Ireland) has advised that a solution of the above problem appears in a joint paper by Hans Liebeck and himself, *Groups of odd order with automorphisms inverting many elements*, J. London Math. Soc. (2), 6 (1973), 215–223. He has also obtained the following generalization: G finite, p least prime dividing $|G|$, $\alpha \in \mathrm{Aut}(G)$, $T_n = \{g \in G \mid g\alpha = g^n\}$, $n \in \mathbb{Z}$; then if $|T_n| > |G|/p$ we have $T_n = G$. See J. London Math. Soc., 11 (1975), 366–368.

Smoothing a Continuous Function

6027 [1975, 409]. *Proposed by Philip Hanser, Columbia University*

Let f be a continuous real function on R, the reals. Must there exist a strictly increasing function $g : R \to R$ such that $g \circ f$ is everywhere differentiable?

Solution by Stuart Jay Sidney, University of Connecticut. No. In fact, if f is continuous but nowhere differentiable while g is strictly increasing, then $g \circ f$ cannot be differentiable on any interval; if in addition g is continuous, then the same conclusion obtains for $f \circ g$.

To see the first of these assertions (the proof of the second is similar), let I be any open interval and suppose $g \circ f$ is differentiable throughout I. Then $f(I) = J$ and $g(J) = (g \circ f)(I) = K$ are non-degenerate intervals. $G = g \mid J$ maps J monotonically onto K, so $G^{-1} : K \to J$ maps K monotonically onto J, so is differentiable almost everywhere on K. Choose y in the interior of K at which G^{-1} is differentiable and choose $x \in I$ such that $(g \circ f)(x) = y$. Since $f = G^{-1} \circ (g \circ f)$ on I, the chain rule tells us that f is differentiable at x, a contradiction.

Also solved by Thad Dankel, Jr., Roy Davies (England), M. B. Gregory, William Margulies, David Ritter, Gideon Schwarz, Oto Strauch (Czechoslovakia), Peter Ungar, and Joseph Ullman.

Identically Distributed Random Variables

6030 [1975, 528]. *Proposed by David Griffeath, Cornell University*

Let X and Y be jointly distributed real random variables. Consider the conjecture: If X, Y, $X + Y$, and $X - Y$ are all identically distributed, then $X = 0$ almost surely.

Prove or disprove the conjecture in the following cases:

(a) if X is square-integrable;

(b) if X is integrable;

(c) in general.

Solution by Peter Hooper, Queen's University, Kingston, Ontario.

(a) If X is square-integrable we have

$$EY^2 = E(X + Y)^2 = E(X - Y)^2$$

which expands to

$$EY^2 = EX^2 + EY^2 + 2EXY = EX^2 + EY^2 - 2EXY.$$

Here the second equality gives $EXY = 0$, and the first equality then gives $EX^2 = 0$, showing that $P(X = 0) = 1$.

(b) Here $E|X| = E|Y| = E|X + Y| = E|X - Y|$. Now $|X + Y| + |X - Y| \geqq |X + Y + X - Y| = 2|X|$, whence $|X + Y| + |X - Y| - 2|X| \geqq 0$. But $E\{(|X + Y| + |X - Y| - 2|X|)\} = 0$. So $|X + Y| + |X - Y| - 2|X| = 0$ with probability 1. Similarly $|X + Y| + |X - Y| - 2|Y| = 0$ with probability 1. So $|X| = |Y|$ with probability 1.

The joint distribution of X, Y is entirely concentrated on the two lines $Y = X$ and $Y = -X$ through $(0, 0)$. The amounts of probability on these two lines are $P(X - Y = 0)$ and $P(X + Y = 0)$ respectively, while $P(X = 0)$ is the amount of probability on their intersection point. Since these three probabilities must be equal, all the probability must be on the intersection point: $P(X, Y = 0, 0) = 1$.

(c) Counterexample: Let U and V be independently distributed each with Cauchy density f given by $f(x) = 1/\pi(1 + x^2)$. Define $X = (U + V)/2$, $Y = (U - V)/2$. Then X, Y, $X + Y = U$, and $X - Y = V$ are each distributed with Cauchy density f.

Also solved by Thad Dankel, Jr., G. Fourt (France), Ellen Hertz, S. Kalikow, O. P. Lossers (Netherlands), G. S. Rogers & D. L. Young, and the Stanford Statistics Problem Solving Group.

A Random Walk Application

6031 [1975, 528]. *Proposed by I. I. Kotlarski, Oklahoma State University*

Let φ be a periodic function on $\mathbb{R}$ with period 2π, given by

$$\varphi(t)=1-\sqrt{\frac{|t|}{\pi}\left(2-\frac{|t|}{\pi}\right)},\qquad t\in[-\pi,\pi]. \tag{1}$$

Prove that φ is a characteristic function of a real random variable X, and find its probability structure.

Let $X_1, X_2, \ldots, X_n, \ldots$ be independent identically distributed random variables all distributed according to the characteristic function (1). Denote

$$Y_n=(X_1+X_2+\cdots+X_n)/n,\ Z_n=(X_1+X_2+\cdots+X_n)n^2.$$

Show that Y_n does not have a limiting distribution while the limit distribution of Z_n is the stable distribution with exponent $\frac{1}{2}$.

Solution by the Stanford Statistics Problem Solving Group. The periodic function f on $\mathbb{R}$ with period 2π given by

$$f(t)=1-\frac{|t|}{\pi},\qquad t\in[-\pi,\pi]$$

is the characteristic function of an integer-valued random variable W where

$$P(W=0)=\tfrac{1}{2},\ P(W=n)=\frac{1-\cos n\pi}{(n\pi)^2},\qquad n\neq 0.$$

(See Feller, *Probability Theory and its Applications*, 2nd Ed., Vol. II, p. 506.) Let N be the distribution of the first return to 0 in a random walk on the line with a fair coin. Then N has the generating function $h(s)=1-(1-s^2)^{1/2}$. If $\{W_i\}_{i=1}^{\infty}$ are independent copies of W then the random sum $\Sigma_{i=1}^{N} W_i$ has characteristic function $h(f(t))=\varphi(t)$.

To show convergence of Z_n we may prove that

$$\lim_{n\to\infty}\left(\varphi\left(\frac{t}{n^2}\right)\right)^n=\lim_{n\to\infty}\left[1-\sqrt{\frac{|t|}{\pi n^2}\left(2-\frac{|t|}{\pi n^2}\right)}\right]^n$$

$$=\lim_{n\to\infty}\left[1-\frac{1}{n}\sqrt{\frac{2t}{\pi}}\right]^n=\exp\left(-\left|\frac{2t}{\pi}\right|^{1/2}\right).$$

Thus Z_n converges to a symmetric stable law of index $\frac{1}{2}$. This yields that $Y_n=nZ_n$ does not have a limit as $n\to\infty$.

Also solved by L. E. Clarke (England), O. P. Lossers (Netherlands), G. S. Rogers, and the proposer.

NOTE. The proposer, in his solution, shows that

$$P(X=k)=p_k=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{-ikt}\varphi(t)dt,\qquad k=0,\pm1,\pm2,\ldots;$$

$$p_0=1-\pi/4,\qquad p_k=p_{-k}=\frac{1}{2k}(-1)^{k-1}J_1(k\pi),\qquad k=1,2,\ldots,$$

where J_1 is the Bessel function of order 1.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN
with the assistance of the mathematics departments of St. Olaf and Carleton Colleges
COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, CARLETON COLLEGE

We invite readers to submit reviews of significant recent college-level mathematics books. We especially encourage reviews based on classroom use, or comparative reviews of several related books. Reviews should ordinarily not exceed two pages (per book) typed double spaced. Manuscripts of reviews as well as books submitted for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057.

A First Course in Linear Algebra. By Raymond A. Beauregard and John B. Fraleigh. Houghton Mifflin, Boston, Massachusetts, 1973. xiii + 410 pp. \$14.50. (Telegraphic Review, April 1973.)

At Auburn University at Montgomery, linear algebra is one of three sophomore level courses which can be taken in any order; the others are multivariable calculus and differential equations. The course is taken primarily by math majors and engineering students. Thus the text must be accessible to students with one year of calculus, must contain enough theory for the math majors, and must give a computational and conceptual background for those requiring serious applications. The book is divided into five parts varying in pace and sophistication, so I will discuss each part separately.

The book begins with a leisurely (55 page) development of geometry in R^n, beginning with geometric vectors in the plane. It covers the usual topics, emphasizing angles between vectors and equations of lines and hyperplanes. The presentation is clear but laborious. The second part deals with matrices and linear equations. Three chapters cover matrix algebra, solutions of systems of linear equations by row reduction, and matrix inversion. There are many references to Part One, and the pace is about the same. There is also an extremely concise section on the Jacobian matrix of a differentiable map, including a matrix statement of the Chain Rule.

Part Three, entitled "Optional Topics in Abstract Algebra", presents a concise (75 page) development of the idea of a vector space starting with semigroups; the level is that of an abstract algebra course. I doubt many classes would have either the background or the time to cover this material.

The book returns to more conventional material with the part which is aptly titled "Fundamentals of Linear Algebra." This section is independent of Part Three, except for a few footnotes. It begins with real vector spaces and continues through subspaces, independence, dimension, coordinates, linear maps, the matrix of a linear map, change of basis, rank, eigenvalues, and determinants (defined by minors, with permutations in an optional chapter). The level is somewhat above average for a sophomore course, but the presentation is excellent: mathematically honest, with nice proofs and thorough explanations. There are exceptionally good conceptual treatments of several subjects: e.g., an elegant treatment of change of basis using the relationship between composition of linear maps and matrix multiplication; an initial discussion of computation of eigenvalues by attempting to invert $A\text{-}\lambda I$; and a chapter on the equation $\phi(x) = b$ (where ϕ is a linear map) which interprets the results in terms of both linear and differential equations. Unfortunately, the applications are less successful. Cramer's Rule appears, but without motivation. There are good applications to linear differential equations; also a section using diagonalization to solve systems of such equations, but with no indication of where these systems occur.

The final part contains discussions of Jordan form, inner product spaces, Gram-Schmidt, unitary maps, and volume change in multiple integrals. The level is the same as in the preceding part. My one quarter course barely began this section; semester courses could include several of these topics.

The exercise sets are adequate, although short on computational problems; but Schaum's *Outline of Linear Algebra* follows a similar development and contains many useful supplementary problems.

Answers are given to odd-numbered problems, even proofs. We found only a few misprints, all obvious.

In summary, I was pleased with the text and so were most of the students. I doubt anyone would follow it precisely, and motivation for the applications should be added. But the development of the theory is excellent, and the authors' expository skill helps students understand the purpose, as well as the mechanics, of the computations.

CHESTER PALMER, Auburn University at Montgomery

Completeness, Compactness, and Undecidability. By Alfred B. Manaster. Prentice-Hall, Englewood Cliffs, New Jersey, 1975. vi + 160pp., $10.25. (Telegraphic Review, June-July 1975.)

This is a new text in mathematical logic, and it has much to recommend it. I used it in a three semester hour course in mathematical logic for advanced undergraduates and beginning graduate students, and plan to use it again.

The three chapters of the text deal with propositional calculus, predicate calculus, and undecidability of predicate calculus and arithmetic. Formal systems are presented using Gentzen sequent logic, that is, theorems are stated in the form $\alpha \to \beta$ where α and β are finite strings of wffs, and the interpretation of such a theorem is that the conjunction of the wffs in α implies the disjunction of the wffs in β. The only axioms are sequents of the form $A \to A$ where A is a statement letter or atomic wff. Rules of inference are presented as tree diagrams, and proofs are constructed as tree diagrams. By the way, there are remarkably few typographical errors considering the complexity of the printing.

There are a number of positive features to this text. First of all, it is superbly organized and structured; it "hangs together" beautifully. The author has resisted the temptation to introduce extraneous material or interesting sidelights. All moves smoothly to a satisfactory conclusion, leaving one with the feeling he has just finished wrapping a tidy and attractive package with exactly the right amount of paper and string. One consequence is that the text is short (150 pages) and can be completely covered in one semester. There are some excellent exercises, some routine ones which almost all the students can do, and a few quite difficult ones. Especially noteworthy are those exercises which demand use of (and understanding of) some of the proof techniques used in proving the metatheorems of the text material. Because this course is also used as an elective in our graduate computer science degree program, I found additional plusses to be the use of Turing machines and the halting problem in the proof of the undecidability of predicate calculus and arithmetic, and the use of tree diagrams.

Of course there are many topics omitted from this book which some people might feel ought to appear. This can be said, I suppose, about any text. Aside from that, there are some disadvantages to this particular book. The language used is very concise and the notation can get a bit complicated. The students found it difficult, although not impossible, to read. Perhaps I should say that better students enjoyed it but it required some help on my part for the material to be accessible to the average student. There certainly could be more example problems for the construction of proofs. There is a particular lack of examples illustrating mathematical formalization of English language sentences, such as those which appear in Suppes, *Introduction to Logic*, Van Nostrand, 1957. The instructor can easily supply such examples in the classroom, however.

In summary, this is a beautifully written text which I would certainly recommend for consideration.

JUDITH L. GERSTING, Indiana University — Purdue University at Indianapolis

TELEGRAPHIC REVIEWS

Telegraphic reviews are designed to give prompt notice of new books with sufficient information to assist our readers in deciding whether to order an examination copy or to suggest library purchase. Possible uses are indicated as follows:

T = textbook P = professional reading
S = supplementary reading L = undergraduate library purchase
13 to 18 = freshman to second year graduate level usage
1 to 4 = appropriate time in semesters to cover text

Asterisks (*) or question marks (?) denote special positive or negative emphasis, respectively. Publishers are denoted by standard abbreviations; complete addresses may be found in *Books in Print*.

GENERAL, S*, L*, *The Incredible Dr. Matrix*. Martin Gardner. Scribners, 1976, 256 pp, $8.95. All of the 18 columns through 1975 that Gardner has written for *Scientific American* on the whimsical Dr. Matrix, so that this book supplants *The Numerology of Dr. Matrix* (Simon & Schuster, 1967), which reprinted 8 columns. Commentary on the earlier columns has been updated, though gone is the "Index of Numbers Under 1,000." PJC

GENERAL, S(13), L, *Playing with Infinity, Mathematical Explorations and Excursions*. Rózsa Péter. Trans: Z.P. Dienes. Dover, 1976, xiii + 268 pp, $3 (P). A popularization of mathematics touching on number theory, Galois theory, geometry, calculus, symbolic logic, Gödel's undecidability theory and much more. Written in Hungarian in 1943, it was officially published in Budapest first in 1957; this edition is an unaltered reprint of the first English translation published in England in 1961. LAS

GENERAL, L*, *Van Nostrand's Scientific Encyclopedia, Fifth Edition*. Ed: Douglas M. Considine. Van N-Rein, 1976, xi + 2370 pp, $67.50. A thorough revision of the 1968 Fourth Edition; nearly 80% of the material in this volume is new. Mathematics accounts for less than 5% of the volume; most of the 500 mathematics entries are to basic definitions (e.g., matrix, group, integral). A handy (albeit voluminous) reference for basic scientific information, but only of incidental usefulness for mathematics; one could do as well with nearly any other standard handbook. LAS

GENERAL, L, *Islamic Patterns, An Analytic and Cosmological Approach*. Keith Critchlow. Schocken Books, 1976, 192 pp, $24.95. "...shows how the art of Islam is inseparable from the science of mystical mathematics associated with the Pythagorean tradition." "The thesis of the present book is that these self-evident mathematical patterns with their esoteric philosophical values became the invisible foundation upon which the 'art' was built." The author takes note of regular and semi-regular tessellations but makes no mention of the instances of two-dimensional crystallographic patterns in Islamic art. PJC

PRECALCULUS, T(13), *Algebra and Trigonometry*. Herbert A. Hollister. Har-Row, 1976, vii + 470 pp, $12.95. Assuming geometry and elementary algebra, this text covers equations and inequalities, graphs and functions, logarithms, analytic and numerical trigonometry, complex numbers and sequences. Answers to odd-numbered problems and index. PJM

HISTORY, S*, P*, L***, *Courant in Göttingen and New York*. Constance Reid. Springer-Verlag, 1976, 314 pp, $12.80. Courant, a "Hilbert in the spirit of Klein", realized Klein's dream by establishing the Mathematical Institute at Göttingen and then by extraordinary perseverance and conniving transplanted it in spirit as well as in substance to New York University. This biography, consisting largely of extended frank quotations from interviews, letters, and diaries, focuses more on Courant's life than on his mathematics. Stories about wartime incidents, never before recorded, provide brutal insight into the personal and professional tensions of that era. An "incurable optimist", Courant pursued his ambitions with an unexpectedly effective combination of insight and indecisiveness. Reid attributes Courant's success more to his political skills than to his mathematical leadership: "He understood human beings and life." LAS

HISTORY, P, L*, *Men and Institutions in American Mathematics*. Ed: J. Dalton Tarwater, John T. White, John D. Miller. Texas Tech Pr, 1976, 136 pp, $5 (P). Ten papers from the May 1973 bicentennial conference at Texas Tech University; authors include Marshall Stone, Garrett Birkhoff, Salomon Bochner, Dirk Struik. LAS

FOUNDATIONS, P, *Boole's Logic and Probability*. Theodore Hailperin. Stud. in Logic and Found. of Math., V. 85. North-Holland, 1976, x + 245 pp, $24.50. Despite the historical importance of Boole's work in logic, there has been no detailed explanation of why the algebraic methods he used are successful in logic. (His methods differed from the Boolean algebras they suggested to later workers.) Boole's ideas on the foundations of probability are neither well-known nor well-understood. This book explains what exactly Boole did, by using new contemporary notions. The investigation is explanatory in spirit, not historical. PJC

FOUNDATIONS, P, *Lecture Notes in Mathematics-537: Set Theory and Hierarchy Theory, A Memorial Tribute to Andrzej Mostowski*. Ed: W. Marek, M. Srebrny, A. Zarach. Springer-Verlag, 1976, xiii + 345 pp, $12.30 (P). Proceedings of the Second Conference on Set Theory and Hierarchy Theory, Bierutowice, Poland, September 1975. Includes *Curriculum vitae* of Mostowski, together with a bibliography of his works. Notable papers in the volume include A. Sochor's "The alternative set theory" which describes a consistent set theory created by P. Vopěnka in which there are only two infinite cardinals but which seems viable as a vehicle for mathematics. Browsers will enjoy deciphering Devlin's Axiom (p. 67), whose author-namesake claims "is destined to be the world's most useless axiom of set theory." PJC

FOUNDATIONS, P, L. *The Concepts of Space and Time: Their Structure and Their Development.* Ed: Milič Čapek. Boston Stud. in Philo. of Sci., V. 22. Reidel, 1976, lvii + 570 pp, $58. An anthology of 75 reprinted selections by philosophers, physicists and mathematicians divided into three parts: classical (pre-relativistic) concepts of space, of time, and modern (post-relativistic) concepts of space-time. LAS

COMBINATORICS, T(17), P*, L. *Matroid Theory.* D.J.A. Welsh. Acad Pr, 1976, xi + 433 pp, $38. An up-to-date introduction to and survey of the field. The theme is the unifying role of matroids in combinatorial theory. The author covers the basics of the theory as well as the connection between matroids and lattices, graphs, transversal theory, covering problems, coding, geometric structures and algebraic structures. Included are a number of exercises, useful notes on each chapter, and an extensive bibliography. An important book. SG

NUMBER THEORY, S(15-16). *Zahlentheorie.* Georg Johann Rieger. Vandenhoeck & Ruprecht, 1976, 219 pp, DM 39 (P). Beginning with unique factorization and congruences, the author discusses quadratic reciprocity, number theoretic functions, sieve methods, p-adic numbers, diophantine approximation, continued fractions, transcendental numbers, the prime number theorem, Dirichlet's theorem, and Waring's problem. No exercises. SG

LINEAR ALGEBRA, T(15-16: 1), S, L. *Matrix Theory and its Applications, Selected Topics.* N.J. Pullman. Pure and Appl. Math., V. 35. Dekker, 1976, vi + 240 pp, $21.75. A readable, theoretical treatment of advanced topics selected for their applicability: non-negative matrices, Perron's theorem, stochastic matrices, differential equations, stability, location of eigenvalues, Gersgorian's theorem. An introductory chapter reviews and extends prerequisite notions from elementary linear algebra. A good option for an elective sequel to the standard introductory linear algebra course.LAS

ALGEBRA, P. *Geometric Algebra over Local Rings.* Bernard R. McDonald. Pure and Appl. Math., V. 36. Dekker, 1976, xii + 421 pp, $29.50. An excellent survey of geometric algebra over commutative rings with special emphasis on local rings. Topics: local rings and their spaces, normal subgroups and automorphisms of the general linear group, the symplectic group, the orthogonal group; bilinear forms and isometry groups. Includes a large bibliography. A valuable reference. SG

ALGEBRA, T*(15-16: 2), S*, P, L. *Modern Algebra with Applications.* William J. Gilbert. Wiley, 1976, xii + 348 pp, $21.95. Boolean algebra and its application to switching circuits, groups and symmetry (e.g., Polya-Burnside enumeration theorem), monoids and finite state machines, rings and coding, fields and the construction of orthogonal Latin squares. A rather condensed account of the group-ring-field theory contained in a standard one term beginning abstract algebra course can be extracted from the text--stress in the exercises is on specific instances and illustrations of the theory. LCL

CALCULUS, T(13: 2). *Calculus and Its Applications.* Larry J. Goldstein, David C. Lay, David I. Schneider. P-H, 1977, xiv + 530 pp, $13.95. An informal treatment assuming a "minimum of prerequisite knowledge" with an unusually extensive collection of realistic applications to the biological and behavioral sciences. Even the product and chain rules are banished from the first half of the course to make room for an overview of all essential ideas of polynomial and exponential calculus. The second half is devoted to more advanced techniques, including functions of several variables, trigonometric and logarithmic functions, and various differential equations. More substantial than most books of its genre, its complete absence of theory may nevertheless be a liability for those students who continue their study of mathematics. LAS

CALCULUS, T(13: 2, 3). *Calculus: An Intuitive and Physical Approach, Second Edition.* Morris Kline. Wiley, 1977, xvi + 943 pp, $18.95. Many improvements over the first edition. Good format. Additions includes applications to economics, a chapter on differential equations, and a section on numerical methods. LLK

DIFFERENTIAL EQUATIONS, T(16-17: 1). *Introduction to Partial Differential Equations with Applications.* E.C. Zachmanoglou, Dale W. Thoe. Williams & Wilkins, 1976, x + 405 pp, $15.95. Introductory, with many examples and applications. Integral curves and surfaces, quasi-linear and linear first order equations, Cauchy-Kovalevsky theorem, theory of linear equations, Laplace's, wave, heat equations. Includes review of needed material from calculus and ordinary differential equations, and a guide to further study. From a course at Purdue. DFA

DIFFERENTIAL EQUATIONS, P. *Function Theoretic Methods in Differential Equations.* Ed: R.P. Gilbert, R.J. Weinacht. Pitman, 1976, 309 pp, $9.50 (P). A collection of 19 papers covering three broad topics: generalizations of analytic function theory, boundary value problems, integral equations. SG

FUNCTIONAL ANALYSIS, P. *Charting the Operator Terrain.* John Ernest. Memoirs No. 171. AMS, 1976, iii + 207 pp, $8.80 (P). From author's abstract: "The purpose of this memoir is to offer a cartographic procedure for bringing some organizational sense to the prodigious task of exploring and describing the vast and varied terrain of bounded operators on a separable Hilbert space. ...Our cartographic system is essentially an adaptation to single operators of John von Neumann's classification and reduction theory." I-CH

OPTIMIZATION, T*(15-17: 1), S*, L*. *Methods of Optimization.* G.R. Walsh. Wiley, 1975, x + 200 pp, $19.95; $9.95 (P). Contains almost all the fundamental ideas in modern optimization theory. Presents important methods in a straightforward and systematic way. Exercise problems, often small-scaled (involving few variables), are intended to illustrate the theory in a simple way. An excellent textbook for a course in 'static optimization' (as opposed to optimal control theory). I-CH

OPTIMIZATION, P, *Machine Scheduling Problems: Classification, Complexity and Computations.* A.H.G. Rinnooy Kan. Martinus Nijhoff, 1976, ix + 180 pp, Dfl. 50 (P). Based on the author's doctoral dissertation. Classifies machine scheduling problems, e.g., as "flow-shop" or "job-shop." Discusses solution methods, e.g., combinatorial, branch and bound, dynamic programming. Obtains bounds on the complexity of some types of problems. Prerequisites are linear programming and graph theory. Bibliography. RWN

COMPUTER SCIENCE, *Finite State Fantasies.* Rich Didday. Matrix Pub, 1976, 48 pp, $2.25 (P). A collection of short cartoon stories with the computer as leading character in the style of Zap Comix or Swann and Johnson's *Prof. E. McSquared's Original, Fantastic & Highly Edifying Calculus Primer* (TR, June-July 1975). PJC

COMPUTER SCIENCE, T(14), S, L, *Minicomputer Systems: Organization and Programming (PDP-11).* Richard H. Eckhouse, Jr. P-H, 1975, xviii + 343 pp, $15.95. Computer organization and assembly language (PAL-11) programming. Includes addressing schemes, programming techniques, data structures, I/O, system software and operating systems. Several appendices. Exercises. RWN

COMPUTER SCIENCE, S(16-17), P, L, *The Design of an Optimizing Compiler.* William Wulf, *et al.* Am Elsevier, 1975, x + 165 pp, $8.75 (P); $13.50. The design and implementation of a compiler to produce very efficient object code. Although concerned with a particular language (BLISS) on a particular machine (the PDP-11) a general viewpoint is maintained. The effects of the language design on the compiler are also studied. Appendices on BLISS and the PDP-11. RWN

COMPUTER SCIENCE, T(13: 1), S, *Non-Technical FORTRAN.* Thomas Worth. P-H, 1976, 356 pp, $9.25 (P). Only mathematical prerequisites for this text are one year of high school algebra and rudiments of trigonometry. Complete explanations of the various language features. Many examples (including worked-out problems) and exercises. Extensive coverage of timesharing FORTRAN. An extensive, convenient index. RJA

COMPUTER SCIENCE, T(13: 1), S, *COBOL for Small and Medium-sized Computers.* Asad S.O. Khailany, Claude V. Duplissey. HM, 1976, xiii + 360 pp, $9.95 (P). Text on COBOL as it is implemented on the IBM-1130; however, all the essentials of COBOL programming are included. In addition, mentions differences for the IBM System 3, the Burroughs 1700, and the DECSYSTEM-10 in appendices. Chapter summaries, review questions with selected answers, and suggested projects. References. Index. RJA

COMPUTER SCIENCE, T(13: 1), S, *Minicomputers, Structure & Programming.* T.G. Lewis, J.W. Doerr. Hayden, 1976, 282 pp, $12.95. An introductory text for a machine organization and assembly language programming course. Divided into 3 parts: number conversion and codes, minicomputer organization, and minicomputer software. Explanations are clear with many detailed examples. Tables and diagrams which highlight the important points are included throughout. Problems. Chapter references. Index. RJA

COMPUTER SCIENCE, T(13: 1), S, *Advanced Basic: Applications and Problems.* James S. Coan. Hayden, 1976, 184 pp, $6.95 (P); $8.95. Begins with a review chapter on the elementary features of the language. Contains chapters on strings, files, and extended features of BASIC. Last two-thirds of text deals with applications programming in coordinate geometry, statistics, simulation, game playing, and in uses of polynomials, sequences, series and matrices in programming. Problems and selected answers. Appendices. Bibliography. Index. RJA

COMPUTER SCIENCE, T(13: 1), S, *A Primer on PASCAL.* Richard Conway, David Gries, E.C. Zimmerman. Winthrop, 1976, xii + 433 pp, $9.95 (P). An introduction to programming using an elementary subset of PASCAL. PASCAL is an excellent vehicle for teaching good programming habits, from initial design considerations to final coding. Points are amply and thoroughly illustrated in this text. Appendices. Exercises. References. Index. RJA

COMPUTER SCIENCE, S(17-18), P, *Lecture Notes in Computer Science-43: Komplexität von Entscheidungsproblemen.* Ernst Specker, Volker Strassen. Springer-Verlag, 1976, 217 pp, $9.50 (P). Eleven papers given at the 1973-74 Zurich seminar on the complexity of logical and combinatorial decision problems. JD-B

COMPUTER SCIENCE, P, *Real-Time Software.* Infotech Inter, 1976, ix + 880 pp. This is a state of the art report on the subject of the title. Includes sections on architecture, operating systems, teleprocessing, languages, data management, performance, reliability, and applications. Second part of the work consists of papers by individual authors on specific topics in Real-Time Software. Bibliography. Several indexes. RJA

COMPUTER SCIENCE, S*(15-18), P*, L*, *Mariages stables et leurs relations avec d'autres problèmes combinatoires.* Donald E. Knuth. Pr U Montreal (US Distr: Intern. Schol. Book Serv.), 1976, 106 pp, $10. Knuth uses the "stable marriage problem" (designing pairings so that no two individuals who prefer each other are married to others) as an occasion for a seven-lecture introduction to the analysis of algorithms. The result is a beautiful tapestry woven from threads of combinatorics, probability, analysis, and algebra, as well as data structures, control structures, and computational complexity. The last lecture discusses open problems in detail. The author notes that the book does not demand previous experience either in analysis of algorithms or in marriage. High price for typescript with unneeded hard cover. PJC

COMPUTER SCIENCE, T(13: 1), S, L, *Computational Mathematics.* Gideon Zwas, Shlomo Breuer. Univ Pub Projects (Israel), 1975, 230 pp, (P). Part A presumes no calculus and considers compound interest, equation solving, computation of areas, approximating π, solving linear systems and approximation by polynomials. Part B reconsiders and extends these problems using calculus. Exercises. RWN

COMPUTER SCIENCE, T(15). *Principles of Continuous System Simulation.* W.K. Giloi. Teubner, Stuttgart, 1975, 172 pp, (P). Subtitle: Analog, Digital and Hybrid Simulation in a Computer Science Perspective. Lecture notes, software and programming techniques. Mathematical bases. Interplay between analog and digital simulation. Few explicit exercises. RWN

SYSTEMS THEORY, P*, L*. *On Systems Analysis: An Essay Concerning the Limitations of Some Mathematical Methods in the Social, Political, and Biological Sciences.* David Berlinski. MIT Pr, 1976, 186 pp, $15. A serious, thorough, uncompromising attack on the "inflated" theory of systems analysis: there is in systems analysis an "inexpungible craving for generality", an "immemorial intellectual inclination" without which mathematics would be sterile and physics uninteresting. "What separates systems analysis from such serious studies is simply the gaps that inexorably open between the conception and execution of a set of intellectual ambitions." Philosopher Berlinski illustrates these "gaps" with well-documented examples from a wide range of systems literature, substantiating his main theme that the differential equation models of mathematical physics are generically unsuited to the social sciences. LAS

SYSTEMS THEORY, T(16-17: 1, 2), S. *Introduction to Linear Systems Analysis.* George M. Swisher. Matrix Pub, 1976, xv + 724 pp, $22.95. Intended to "bridge the gap from classical controls (transform analysis) to modern systems analysis (state-space techniques)." A prior course in classical linear control theory would be desirable, but not necessary. FLW

SYSTEMS THEORY, T(16-17: 1). *Introduction to Discrete Linear Controls: Theory and Application.* Albert B. Bishop. Acad Pr, 1975, xiii + 378 pp, $29.50. Written for readers interested in conceptual models of various discrete, linear, and time-invariant systems. Assumes no background in control theory nor any particular knowledge of discrete mathematics. The first half of the book describes discrete systems by linear difference equations and their solutions. The second half deals mainly with system-performance evaluations; here knowledge of elementary probability theory is assumed. I-CH

APPLICATIONS, P, L. *Surveys in Applied Mathematics.* Ed: N. Metropolis, S. Orszag, G.-C. Rota. Acad Pr, 1976, xv + 297 pp, $14. A valuable collection of 18 essays by distinguished authors (e.g., C.C. Lin, R. Bott, J.D. Cole, J. Glimm, P. Lax) on current topics in applied mathematics. From a workshop at Los Alamos dedicated to S.M. Ulam. All the articles are reprinted from Volumes 16 and 19 of *Advances in Mathematics.* LAS

APPLICATIONS, P, L. *Analyse numérique des inéquations variationnelles.* R. Glowinski, J.L. Lions, R. Trémolières. Dunod (US Distr: SMPF), 1976. *Tome 1: Théorie générale premières applications,* xii + 268 pp, 180 F; *Tome 2: Applications aux phénomènes stationnaires et d'évolution,* xiv + 290 pp, 210 F. A variational inequality is a constraint on the partial derivatives of a function. These two volumes deal with numerical methods for solutions to such inequalities, and applications in physics, control theory and biology. PJM

APPLICATIONS (BIOLOGY), T*(15-16: 1, 2), P, L*. *Mathematical Bioeconomics: The Optimal Management of Renewable Resources.* Colin W. Clark. Wiley, 1976, xi + 352 pp, $21.95. A systemmatic survey of the developing theory of conservation of productive resources (e.g., fish, forests, orchards), using only basic techniques of calculus. Replete with examples of current interest and sometimes surprising conclusions (e.g., where immediate extermination of a species may be the most "profitable" long-range policy), this monograph would provide an excellent resource for a seminar in modern applied mathematics. LAS

APPLICATIONS (CHEMISTRY), T?(18: 1), P. *Group Theoretical Techniques in Quantum Chemistry.* C.D.H. Chisholm. Acad Pr, 1976, x + 271 pp, $22.25. Written for graduate students studying quantum chemistry--basic concepts and results of quantum mechanics are assumed. Treats finite molecular symmetry groups, permutational symmetry and the symmetric group, spin-free quantum chemistry, theory and representation of continuous groups, plus advanced chemical topics. Detailed proofs are omitted but references are supplied. PJC

APPLICATIONS (ECONOMICS), S(14). *GEM, A General Econometric Matrix Program.* Lucy Joan Slater. Cambridge U Pr, 1976, vii + 86 pp, $5.95 (P). Description of a canned package for manipulating matrices for econometric model building. Based on Fortran. Includes directions and examples on usage, source listings, implementation details, and extensions to data bank maintenance. RWN

APPLICATIONS (ECONOMICS), T(17-18: 1), S, P. *Econometrics.* Peter Schmidt. Statistics, V. 18. Dekker, 1976, vi + 269 pp, $19.50. Supplies the proof for "fundamental econometric results" that are often omitted in standard econometrics texts. Could be used as a supplement to such. FLW

APPLICATIONS (ECONOMICS), T(16-18: 1, 2), S. *Foundations of Econometrics.* Albert Madansky. Amer Elsev, 1976, viii + 266 pp, $13.50 (P). Matrix theory and multivariate statistical analysis and a short chapter on econometric model building. "In practice the book should be used as a second course in econometrics for those students who want to understand the basis for the techniques." FLW

APPLICATIONS (INFORMATION THEORY), P. *Information Theory, New Trends and Open Problems.* Ed: G. Longo. Springer-Verlag, 1975, v + 340 pp, $23 (P). Proceedings of a 1975 summer school. Contains 3 papers on future directions of information theory, 4 papers on "classical" information theory, and 7 papers on coding theory. SG

APPLICATIONS (LINGUISTICS), P. *The Linguistic Theory of Numerals.* James R. Hurford. Cambridge U Pr, 1975, xi + 293 pp, $25. Investigation of the formal structure of how human languages associate arbitrary words with the universal concepts of numbers. The author offers a theoretical framework in the generative grammar tradition (à la Chomsky), derived from a data base of numerals in several languages. Primarily intended for linguists, though both the subject matter and the method are tangentially related to mathematics. PJC

APPLICATIONS (MECHANICS), S*(13-18), L. *Motion Geometry of Mechanisms.* E.A. Dijksman. Cambridge U Pr, 1976, xvi + 288 pp, $22.50. A detailed study of the geometry of how machines (gadgets) move. Readers who enjoyed a spirograph or were intrigued by a plagiograph will find it all here and then some. JAS

APPLICATIONS (MECHANICS), S(18), P. *Discrete Mechanics, A Unified Approach.* Hayrettin Kardestuncer. Springer-Verlag, 1975, 59 pp, $5 (P). Monograph which, first, provides a short introduction to each of three approximate methods (calculus of variations, finite differences, finite elements) in mechanics, then argues for the unification of these methods in considering certain problems. Discursive. Readable. Ample bibliography. From lectures at a UNESCO conference. DFA

APPLICATIONS (MECHANICS), P. *Mathematical Methods of Two-dimensional Elasticity.* A.I. Kalandiya. Trans: M. Konyaeva. MIR, 1975, 351 pp. An English translation of the original 1973 Russian edition. The book's main theme: presentation of new results, some appearing for the first time, in the solution methods of elasticity theory (abundant in problems of plane deformation and plate bending due to the effects of holes, etc.). Mathematics used (and assumed): complex function theory and singular integral equations. I-CH

APPLICATIONS (PHYSICS), P. *Lecture Notes in Physics-50: Group Theoretical Methods in Physics.* Ed: A. Janner, T. Janssen, M. Boon. Springer-Verlag, 1976, xiii + 627 pp, $20.50 (P). Proceedings of the Fourth International Colloquium held at Nijmegen in June 1975. The more than sixty papers are grouped by such topics as: Gauge Groups; Elementary Particles; Geometric Quantization; Broken Symmetries; Coherent States; Atomic &Solid State Physics; Mathematical Physics; and even Education. JAS

APPLICATIONS (PHYSICS), T(18), P. *Lecture Notes in Physics-46: Hermitian and Kählerian Geometry in Relativity.* E.J. Flaherty. Springer-Verlag, 1976, viii + 365 pp, $13.20 (P). An introduction to the differential geometry of complex manifolds, for physicists. From the author's Ph.D. thesis and subsequent research. Prerequisites: elementary knowledge of general relativity, differential geometry, complex analysis. DFA

APPLICATIONS (PHYSICS), P. *Lecture Notes in Mathematics-503: Applications of Methods of Functional Analysis to Problems in Mechanics.* Ed: P. Germain, B. Nayroles. Springer-Verlag, 1976, xix + 531 pp, $17.30 (P). Joint symposium of l'Union Internationale de Mécanique Théorique et Appliquée and l'Union Internationale de Mathématiques held in Marseille, 1975. JAS

APPLICATIONS (PHYSICS), P. *Lecture Notes in Mathematics-522: Short Wave Radiation Problems in Inhomogeneous Media: Asymptotic Solutions.* Clifford O. Bloom, Nicholas D. Kazarinoff. Springer-Verlag, 1976, v + 104 pp, $7.40 (P).

APPLICATIONS (PHYSICS), T(15-16: 1), *A Short Course in Fluid Mechanics.* Thomas J.R. Hughes, Jerrold E. Marsden. Publish or Perish, 1976, v + 162 pp, $7.50 (P). Appropriate for a one term seminar at the junior-senior level, assuming vector calculus and some complex variables. The rigor is selective, omitting more difficult proofs. Enough references and proposed studies make this attractive for the seminar style. TAV

APPLICATIONS (PHYSICS), P. *The Scale Coordinate and Its Geometry, The Quantization of Riemannian Geometry.* William Bender. Exposition Pr, 1975, vi + 60 pp, $10. Establishes a possible formalism for the quantization of Riemannian geometry using scale coordinates. First part defines scale coordinates and establishes a connection between analysis and number theory. Second part uses scale coordinates in the study of differential equations. References. RJA

APPLICATIONS (PHYSICS), P. *Functional Integration and Its Applications.* Ed: A.M. Arthurs. Clarendon Pr, 1975, x + 195 pp, £9.00. Proceedings of an international conference held in April 1974 at Cumberland Lodge, Windsor Great Park, London. JAS

APPLICATIONS (PHYSICS), P. *Sur quelques questions d'analyse, de mécanique et de contrôle optimal.* Jacques Louis Lions. Pr U Montreal (US Distr: Intern. Schol. Book Serv.), 1976, 211 pp, $14. Inequalities involving partial derivatives: solutions and applications to mechanics and optimal control. PJM

APPLICATIONS (PHYSICS), S(15), P. *Elements of Applied Mathematics.* Ya. B. Zeldovich, A.D. Myškis. Trans: George Yankovsky. MIR, 1976, 656 pp. According to the preface, the book is "to provide the reader with methods and information of practical utility, while simplifying in the extreme all formal definitions and proofs." Topics are from calculus, numerical analysis, complex variable theory, linear differential equations, vector analysis, calculus of variations, probability. DFA

APPLICATIONS (PHYSICS), P. *Lecture Notes in Physics-53: Lectures on Geometric Quantization.* D.J. Simms, N.J.J. Woodhouse. Springer-Verlag, 1976, 166 pp, $8.20 (P). Notes based on lectures given by the first author in London in the autumn of 1974. These lectures presented a partial formulation of the relationship between classical and quantum mechanics as a relationship between symplectic manifolds and Hilbert spaces. JAS

Reviewers Whose Initials Appear Above

Richard J. Allen, St. Olaf; David F. Appleyard, Carleton; Paul J. Campbell, St. Olaf; John Dyer-Bennet, Carleton; Steven Galovich, Carleton; Ih-Ching Hsu, St. Olaf; Lorraine L. Keller, St. Olaf; Loren C. Larson, St. Olaf; Pierre J. Malraison, Carleton; R.W. Nau, Carleton; J. Arthur Seebach, Jr., St. Olaf; Lynn A. Steen, St. Olaf; T.A. Vessey, St. Olaf; Frank L. Wolf, Carleton.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, NW, Washington, D. C. 20036. Items must be submitted at least five months before publication can take place.

PERSONAL ITEMS

Illinois Wesleyan University: Dr. Phyllis Graham Parr has been appointed Associate Professor; Professor Evelyn K. Wantland, Chairman of the Department of Mathematics, retired in August 1976; Dr. William K. Smith, New College, Sarasota, has been appointed Professor and Chairman of the Department of Mathematics.

Dr. Geraldine A. Jensen, Western Connecticut State College, has been promoted to Professor.

Associate Professor D. G. Malm, Oakland University, has been promoted to Professor.

Dr. Franz Streit, University of Bern, has been appointed *professeur extraordinaire* of mathematical statistics at the University of Geneva, Switzerland.

Professor Carroll O. Wilde, Naval Postgraduate School, has been appointed Chairman of the Department of Mathematics.

Dr. Ralph L. Jeffery, Acadia University, Wolfville, Nova Scotia, died on December 12, 1975, at the age of 85. He was a member of the Association for forty-eight years.

1978 COLLOQUIUM OF THE NEW ZEALAND MATHEMATICAL SOCIETY

The 1978 Colloquium of the New Zealand Mathematical Society is to be the venue for a joint meeting of the Australian and New Zealand Mathematical Societies; it will take place in Christchurch, New Zealand, from May 15 to May 22, 1978.

The meeting is expected to take on an international flavour, with invitations to be sent to a number of mathematicians from beyond Australasia. For further information, please write to Professor G. M. Petersen, Department of Mathematics, Canterbury University, Christchurch, New Zealand.

ALABAMA SPACE AND ROCKET CENTER

The Alabama Space and Rocket Center in Huntsville, Alabama, has announced its program of field trips for the 1976–77 school year. The Center is an official NASA Visitor Information Center and a showcase for America's achievements in space. It is a non-profit educational institution, owned and operated by the State of Alabama.

Information, rates, and reservation forms are available by contacting the Alabama Space and Rocket Center, Tranquility Base, Huntsville, Alabama 35807, (205) 837-3400.

AACJC CAREER STAFFING CENTER

The American Association of Community and Junior Colleges maintains a Career Staffing Center for its member institutions and those individuals who would like to be considered for staff positions at more than 900 member colleges. Write for details to AACJC Career Staffing Center, P. O. Box 298-A, Alexandria, Virginia 22314.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

APRIL-MAY MEETING OF THE NORTH CENTRAL SECTION

The 1976 spring meeting of the North Central Section of the MAA was held at St. Cloud State University, St. Cloud, on April 30 — May 1, 1976. There were 144 persons in attendance, including 86 MAA members.

At the business meeting some minor revisions of the section by-laws were passed, Executive Committee decisions to establish a Section Newsletter and a Section Summer Seminar were announced, and the following elections were completed: Chairman-Elect — Milton Legg, Moorhead State University, Moorhead, three-year term on the Executive Committee; Secretary-Treasurer — Louis Guillou, Saint Mary's College, Winona, third one-year term; Members-at-Large of the Executive Committee — James Baglio, North Hennepin Community College, Minneapolis, and David Storvick, University of Minnesota, Minneapolis, one-year terms.

The principal speaker was Lynn Steen, Saint Olaf College, Northfield, co-editor of the MAA's MATHEMATICS MAGAZINE. His topic was *Catastrophe Theory.*

Joseph Konhauser, Macalester College, St. Paul, was the invited speaker for the Friday evening session. The title of Professor Konhauser's presentation was *Must it be a Circle*?

In preparing for this meeting, an emphasis had been placed on obtaining student participation, as speakers and observers. Forty-eight students attended; eight students contributed papers.

The Saturday contributed papers included:

Morley's theorem and its converse, by Donald Kleven, University of Minnesota, Minneapolis.

Number patterns in Pascal's triangle, by John Hanenburg, St. Cloud State University, St. Cloud.

Elliptical inversions, by Douglas Braff, University of Minnesota, Duluth.

When is a locally one-to-one mapping globally one-to-one?, by Daniel Johnson, Macalester College, St. Paul.

Ghost: a computerized learning game, by Kevin Johnson and Frank Simmons, University of Minnesota, Duluth.

Intermediate compactifications, by James Hatzenbuhler and Don Mattson, Moorhead State University, Moorhead.

A characterization of the integers among Euclidean domains, by Steve Galovich, Carleton College, Northfield.

Supernormal subgroups, by Pierre Malraison, Carleton College, Northfield.

Methods of optimization: a course that attracts students from several areas, by Mohammed Bahauddin, St. Cloud State University, St. Cloud.

On the n-bug problem, by Roger Kirchner, Carleton College, Northfield.

More on Pythagorean triples, by Mary Ann Dahlquist and Gerald Bergum, South Dakota State University, Brookings.

Project CALC Report, by Richard Järvinen, Saint Mary's College, Winona.

Integral distances in a square, by Gary Shute and Ken Yocum, South Dakota State University, Brookings.

LOUIS GUILLOU, *Secretary-Treasurer*

JUNE MEETING OF THE PACIFIC NORTHWEST SECTION

The 1976 meeting of the Pacific Northwest Section of the MAA was held at Portland State University on June 19, 1976.

The following invited addresses were presented:

Invariant functions on matrices, by R. M. Koch, University of Oregon.

Cognitive style mapping and its relationship to the mathematics lab, by Dean McIntire, Central Oregon Community College.

Some applications of non-standard analysis to number theory, by Raymond Mayer, Read College.

Fancy meeting you here — a college algebra counting problem leads to an unexpected encounter with e, by Eugene Maier, University of Oregon.

At the business meeting, the following officers were elected for 1976–77: Chairman, Howard Zink; Chairman Elect, Duane DeTemple; Vice Chairman for two year colleges, Larry Curnutt; Vice Chairman for four year colleges, Roy Dubisch; Secretary-Treasurer, John Herzog.

JAMES CALVERT, *Secretary Treasurer*

CALENDAR OF FUTURE MEETINGS

Fifty-seventh Summer Meeting, University of Washington, August 14–16, 1977.

Sixty-first Annual Meeting, Atlanta, Georgia, January 6–8, 1978.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN, St. Francis College, Loretto, Pennsylvania, April 22–23, 1977.

FLORIDA, University of South Florida, Tampa, March 4–5, 1977.

ILLINOIS, Chicago Loop College, Chicago, May 6–7, 1977.

INDIANA, Wabash College, Crawfordsville, April 30, 1977.

INTERMOUNTAIN

IOWA, Drake University, Des Moines, April 22–23, 1977.

KANSAS, Tabor College, Hillsboro, April 2, 1977.

KENTUCKY, early April. Deadline for papers 6 wks. bef. mtg.

LOUISIANA-MISSISSIPPI, University of New Orleans, Louisiana, February 25–26, 1977.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA, Saturday before Thanksgiving and last Saturday in April.

METROPOLITAN NEW YORK, Sarah Lawrence College, April 24, 1977.

MICHIGAN, Eastern Michigan University, Ypsilanti, May 6–7, 1977.

MISSOURI, University of Missouri, St. Louis, April 29–30, 1977.

NEBRASKA, Nebraska Wesleyan University, Lincoln, April 15–16, 1977.

NEW JERSEY, early November and early May.

NORTH CENTRAL, North Hennepin Community College, Minneapolis, Minnesota, April 29–30, 1977.

NORTHEASTERN, Saturday after Thanksgiving, and third week in June in odd-numbered years.

NORTHERN CALIFORNIA, San Francisco State University, February 26, 1977.

OHIO, Denison University, Granville, April 15–16, 1977.

OKLAHOMA-ARKANSAS, Oral Roberts University, Tulsa, Oklahoma, April 1–2, 1977.

PACIFIC NORTHWEST, second Saturday in June. Deadline for papers 6 wks. bef. mtg.

PHILADELPHIA, Moravian College, Bethlehem, November 19, 1977.

ROCKY MOUNTAIN, Metropolitan State College, Denver, Colorado, April 29–30, 1977.

SEAWAY, State University College at Buffalo, May 6–7, 1977.

SOUTHEASTERN, University of Alabama, Huntsville, April 1–2, 1977.

SOUTHERN CALIFORNIA, Loyola Marymount University, Los Angeles, March 12, 1977.

SOUTHWESTERN, Phoenix College, Phoenix, Arizona, April 22–23, 1977.

TEXAS, Baylor University, Waco, April 1–2, 1977.

WISCONSIN, University of Wisconsin, Oshkosh, Spring 1977.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Denver, February 20–26, 1977.

AMERICAN MATHEMATICAL ASSOCIATION OF TWO YEAR COLLEGES

AMERICAN MATHEMATICAL SOCIETY, University of Washington, August 15–18, 1977.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION, University of North Dakota, Grand Forks, June 13–16, 1977.

ASSOCIATION FOR SYMBOLIC LOGIC, Chicago Sheraton, Chicago, April 28–29, 1977.

ASSOCIATION FOR SYMBOLIC LOGIC, Chicago Sheraton, Chicago, April 28–29, 1977.

ASSOCIATION FOR WOMEN IN MATHEMATICS

CANADIAN SOCIETY FOR HISTORY AND PHILOSOPHY OF MATHEMATICS/SOCIÉTÉ CANADIENNE D'HISTOIRE ET DE PHILOSOPHIE DES MATHÉMATIQUES, Hamilton, Ontario, June 2, 1977.

FIBONACCI ASSOCIATION

INSTITUTE OF MATHEMATICAL STATISTICS, Seattle, Washington, August 14–18, 1977.

MU ALPHA THETA

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Cincinnati, Ohio, April 20–23, 1977.

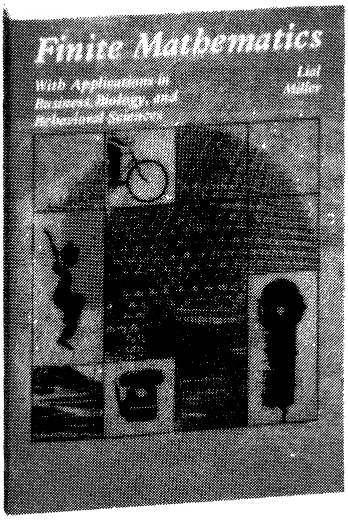
OPERATIONS RESEARCH SOCIETY OF AMERICA, San Francisco Hilton, May 9–11, 1977.

PI MU EPSILON

SCHOOL SCIENCE AND MATHEMATICS ASSOCIATION

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS

Mathematics in



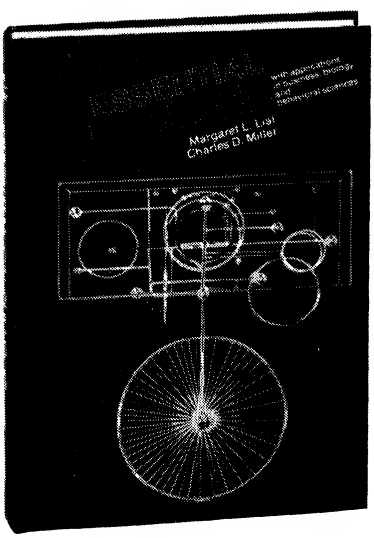
Finite Mathematics

With Applications in Business, Biology, and Behavioral Sciences

Margaret L. Lial / Charles D. Miller

Finite Mathematics is an informal, applied presentation of the mathematical models and tools that students of business, biology, and the behavioral sciences *need* to operate successfully in their chosen fields. Unique features are 17 actual case studies and questions from past CPA examinations, which bring the concepts to life and show students how the mathematics they learn is used in real-world situations. Many examples, word problems, and applications are included throughout. Instructor's Guide with MathLab contains a complete quiz testing program. January 1977, 448 pages, hardbound $12.95

$$\frac{\partial F}{\partial x_1} = 2(-0.7)(0.4 - 0.7x_1 - 0.4x_2) + 2(0.05)(-0.4)(0.15 -$$



Essential Calculus

With Applications in Business, Biology, and Behavioral Sciences

Margaret L. Lial / Charles D. Miller

For a brief, down-to-earth treatment of the fundamentals of calculus, consider the features of *Essential Calculus*: diagnostic pretest and optional algebra review; chapter pretests; 13 case studies showing the real-world applications of calculus to business, biology, and the behavioral sciences; ample drill and word problems throughout; optional sections for longer courses; and an Instructor's Guide with a complete testing program. 1975, 346 pages, hardbound $13.50

COUNT

Houghton Mifflin continues to offer the variety of mathematics texts that teachers count on to meet the diverse needs of students. Here are six new ones:

The Functions of Algebra and Trigonometry
Kenneth Bogart, Dartmouth College
About 500 pages / Instructor's Manual
Just published

Bogart's text is loaded with applications. Almost every topic is introduced and motivated by a word problem. Over 400 illustrations help clarify student understanding as does the informal but accurate style. Every algebraic computation used in calculus is stressed in the book.

"[Bogart] really makes the subject come alive." — D. W. Hall, Michigan State University.

Elementary Algebra by Example
William Brett and Michael Sentlowitz, both of Rockland Community College
About 500 pages / paper / Instructor's Manual / Just published

Brett and Sentlowitz use everyday situations to motivate algebraic concepts and techniques. The steps in solving problems are thoroughly explained in a conversational style. An arithmetic review, over 400 illustrations, numerous problems, glossary, plus chapter and cumulative self-tests further aid learning.

"From the student's standpoint, it might be exactly what they need in order to finally understand algebra. Approach is different and good." — J. L. Amburgey, College of DuPage

Statistics Step-by-Step
Howard B. Christensen, Brigham Young University
About 550 pages / paper / Instructor's Manual / Just published

Hundreds of students at Brigham Young University and at a community college have proven that *Statistics Step-by-Step* works. Christensen's sound pedagogy, generous use of illustrations, both verbal and visual, and step-by-step explanations make his text exceptionally clear. Complete solutions to over 200 problems appear in the text. Sample exams, solutions, and lecture hints for each unit appear in the Instructor's Manual.

Ken Wantling of Montgomery College says, *"The text is the best on explanations that I have seen, has a readable style, covers the appropriate topics for a beginner's statistics course and at an appropriate level."*

Introductory Statistics with Applications
Wayne W. Daniel, Georgia State University
About 450 pages / Instructor's Guide with Solutions / Study Guide / Just published

Daniels provides over 400 applied problems and a wealth of examples in his new elementary statistics text, emphasizing applications to the social and natural sciences and business. Difficult concepts such as sampling distributions and hypothesis testing are clearly explained.

THE CARUS MATHEMATICAL MONOGRAPHS

The Monographs are a series of expository books intended to make topics in pure and applied mathematics accessible to teachers and students of mathematics and also to non-specialists and scientific workers in other fields.

These numbers are currently available:

1. *Calculus of Variations,* by G. A. Bliss.
2. *Analytic Functions of a Complex Variable,* by D. R. Curtiss.
3. *Mathematical Statistics,* by H. L. Rietz.
4. *Projective Geometry,* by J. W. Young.
6. *Fourier Series and Orthogonal Polynomials,* by Dunham Jackson.
7. *Vectors and Matrices,* by C. C. MacDuffee.
8. *Rings and Ideals,* by N. H. McCoy.
9. *The Theory of Algebraic Numbers* (Second edition), by Harry Pollard and Harold G. Diamond.
10. *The Arithmetic Theory of Quadratic Forms,* by B. W. Jones.
11. *Irrational Numbers,* by Ivan Niven.
12. *Statistical Independence in Probability, Analysis and Number Theory,* by Mark Kac.
13. *A Primer of Real Functions* (Second edition), by Ralph P. Boas, Jr.
14. *Combinatorial Mathematics,* by H. J. Ryser.
15. *Noncommutative Rings,* by I. N. Herstein.
16. *Dedekind Sums,* by Hans Rademacher and Emil Grosswald.
17. *The Schwarz Function and its Applications,* by Philip J. Davis.

One copy of each Carus Monograph may be purchased by individual members of the Association for \$5.00 each; additional copies and copies for nonmembers are priced at \$10.00 each. (Orders for under \$10.00 must be accompanied by payment. Prepaid orders will be delivered postage and handling free.)

Orders should be sent to:

MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.
Washington, D.C. 20036

Check these texts for your math course.

A superlative introduction to —

ELEMENTARY LINEAR ALGEBRA, 2nd Ed.

Howard Anton, Drexel University

Your students will appreciate the clarity and style in which this text presents the fundamentals of linear algebra. Abstract concepts are not suddenly thrown at them, but are gradually developed; while all difficult proofs are placed at the end of sections. And a unique treatment is given to the introduction of numerical methods at the level of your students' understanding. Changes in this new edition that make this best-seller even better are:

- New two-color format
- Difficult sections (chapters 4 & 5) have been rewritten in response to users' demands
- Includes applications to Fourier Series, Problems of Approximation, and Systems of Differential Equations.

Examine this book and you'll see why it's been adopted by over 200 colleges and universities.

1977 approx. 376 pp. $13.95

APPLICATIONS OF LINEAR ALGEBRA

Chris Rorres, Drexel University

In this useful supplement, your students find realistic applications, arranged in order of difficulty, with solutions to all exercises. It also includes three chapters on Linear Programming. An ideal supplement to Anton or any other introductory linear algebra text.

1976 approx. 200 pp.
$4.95 (tent.) paper

Essential for the aspiring teacher —

ELEMENTARY MATHEMATICS

Donald F. Devine & **Jerome E. Kaufmann,**
both of Western Illinois University

Here's the right book for showing your students how to teach mathematics to elementary school children. The authors introduce concepts at an intuitive level and present definitions and properties while stressing the "reasoning process." The conversational style of writing gets your students involved as they progress through the logical sequence of topics and gain a knowledge of the foundations of mathematics. Other features that make this book outstanding are:

- Thorough and lucid treatment of the metric system
- Exemplary chapter on algorithms (lauded by reviewers)
- Graded problem sets (with exercises using a hand-held calculator)

Also available—TEACHER'S MANUAL containing solutions to all problems and comments about why certain problems are included.

1976 525 pp. $13.95

To be considered for complimentary examination copies, write to Art Beck, Dept. A 2223-12. Please include course name, enrollment, and title of present text.

JOHN WILEY & SONS, Inc.
605 Third Avenue
New York, N.Y. 10016
In Canada: 22 Worcester Road, Rexdale, Ontario

Prices subject to change without notice A 2223-12

THE MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.

THE AMERICAN
MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

AN OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 84 NUMBER 3

CODEN: AMMYAE

CONTENTS

(*Continued on inside cover*)

MARCH 1977

NOTICE TO AUTHORS

Research papers per se are unsuitable; see statement of policy (Vol. 81, p. 1).

Please follow the format in current issues of the MONTHLY. Manuscripts must be legibly typewritten or reproduced from typewritten copy, double spaced with wide margins. Three copies should be submitted to the appropriate editor and one kept by the author as protection against loss. The author's full address *must* appear at the end of the manuscript.

Backlog: Main Articles 17 months, Math. Notes 26 months, Research Problems 19 months, Classroom Notes 24 months, Math. Education 19 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to R. P. BOAS, Department of Mathematics, Northwestern University, Evanston, IL 60201; NOTES, etc.: to the corresponding Associate Editor; REPRINT PERMISSION: to LEONARD GILLMAN, Mathematical Association of America, University of Texas, Austin, Texas 78712 (see also the copyright notice below); ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N.Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N. W., Washington, D.C. 20036; BACK ISSUES: Contact P. and H. Bliss Co., Middletown, CT 06457.

Annual dues for members of the Association, including a subscription to the American Mathematical Monthly, are \$18.00 for each of the first two years of membership and \$22.00 thereafter. Student Membership is available with annual dues of \$12.00. For nonmembers the subscription price is \$25.00.

PUBLISHED BY THE ASSOCIATION at Washington, D.C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D.C., and additional mailing offices.

DIAMETERS OF ARCS AND THE GERRYMANDERING PROBLEM

JOHN C. OXTOBY

1. Problem and results. Can a city (or a state) be gerrymandered so as to effect any prescribed partition of the residences of the voters? What if no district (connected open set) is allowed to extend more than a given distance outside a minimal square, or circle, that encloses its places of residence, or to have a diameter appreciably exceeding the largest distance between two of these points, and what if the city itself has an irregular boundary? The purpose of this paper is to establish an extremal property of the diameters of plane arcs, and to prove a theorem which incidentally answers these questions although its real motivation lies in its applicability to certain problems in analysis.

Given any partition $F_1, \ldots, F_m$ of a finite subset F of the plane it is easy to find disjoint regions containing the sets F_j; for instance, by constructing disjoint polygonal arcs A_j through the points of each of the sets and then taking sufficiently small neighborhoods of these arcs. The regions R_j so constructed can be made to lie within a given open set U if and only if the arcs A_j can be constructed within U, and they can be made to have nearly the same diameters as these arcs. Hence the problem of effecting a prescribed partition of F by means of disjoint regions, of restricted diameter or within prescribed neighborhoods of the sets F_j, is essentially equivalent to that of constructing disjoint arcs subject to similar restrictions.

Such regions can be useful in various problems. For instance, it may be required to extend a given permutation P of the points of a finite set F to an automorphism T of the plane E^2 in such a way that the norms of the iterates of T and T^{-1} do not exceed a certain bound B. If we partition F into its orbits under P, enclose these in disjoint regions as above, and effect the given (cyclic) permutation on each orbit by means of an automorphism equal to the identity outside all of the regions, then we can take for B the maximum of the diameters of the regions R_j. To restrict B it is therefore sufficient to restrict the diameters of the regions, or equivalently, the diameters of the arcs used in their construction. Thus we are led to pose the following question:

For what values of c (if any) is it true that for every finite family of disjoint finite sets F_j with $\operatorname{diam} F_j > 0$ *there exist disjoint polygonal arcs A_j such that $F_j \subset A_j$ and* $\operatorname{diam} A_j \leqq c \operatorname{diam} F_j$ *for all j?*

Let us remark first of all that in dimension $n > 2$ there is no problem. In that case the points of each set F_j can be joined in any order by broken line segments chosen so as to form disjoint arcs A_j with $\operatorname{diam} A_j = \operatorname{diam} F_j$. Thus in E^n, for $n > 2$, we can even take $c = 1$. But in the plane, if we construct the arcs successively, it may happen that to join two nearby points we are sometimes forced to go far out of the way in order to avoid the arcs previously chosen. Can we forestall such an occurrence, or overcome the difficulty whenever it arises? The following simple example shows that in the case $n = 2$ the number c cannot be taken as small as $2/\sqrt{3}$ (≈ 1.15). The implication is that it may be necessary for the diameter of at least one arc (or region) to exceed the diameter of the corresponding set by about 15%.

Example 1. Let F_1 and F_2 be the vertices of two equilateral triangles inscribed in a circle so as to form a regular hexagon. Let A_1 and A_2 be disjoint arcs containing F_1 and F_2, respectively, and suppose that the diameter of A_1 does not exceed that of the circle. Then A_1 contains no point of the outward ray from any point of F_2. Let A denote the union of A_1 and the three outward rays from the points of F_1. The closed set A divides the plane into regions, three of which are unbounded, and each point of F_2 belongs to a different one of these components. Hence A_2 must intersect the outward ray from one of the points of F_1 (in fact, it must meet at least two such rays). Since the opposite point of the circle also belongs to A_2, the diameter of A_2 must exceed that of the circle, which is $2/\sqrt{3}$ times the diameter of either F_1 or F_2.

This example is critical; we shall see that c can be assigned any value greater than $2/\sqrt{3}$.

Can we require that the arcs be constructed within prescribed neighborhoods of the given sets, say convex neighborhoods? That this is not generally possible, even for rectangular intervals, is shown by the following example.

Example 2. Let R_1 and R_2 be two open rectangles forming a cross. Let F_j $(j = 1, 2)$ contain a point of each component of $R_j - \mathrm{cl}(R_1 \cap R_2)$. Then it is obviously impossible to join the points of F_1 and F_2 by disjoint arcs within R_1 and R_2, respectively.

Similar considerations apply to square neighborhoods in different orientations. But what if the prescribed neighborhoods are circular disks, or square intervals? Then we shall show that the construction is always possible. Even more: it is sufficient that there exist an open set U such that each of the prescribed neighborhoods is a component of a set of the form $U \cap D_j$, where each D_j is an open disk (or a square interval). Our basic result can be formulated as follows.

THEOREM 1. *Let U be an open subset of the plane, let $F_1, \ldots, F_m$ be disjoint finite subsets of U, and let $D_1, \ldots, D_m$ be open disks. There exist disjoint polygonal arcs $A_1, \ldots, A_m$ such that $F_j \subset A_j \subset D_j \cap U$ $(j = 1, \ldots, m)$ if and only if, for each j, F_j is contained in a single component of $D_j \cap U$.*

The condition is obviously necessary (and sufficient) for the existence of each arc separately. The theorem asserts that if each individual arc can be constructed within U and the associated disk, then the arcs can be constructed disjointly.

Before proving the theorem let us see how it settles the question of the possible values of c. For this purpose we need in addition the following well-known theorem, due to H. W. E. Jung.

THEOREM 2 (Jung). *Let F be a bounded closed subset of the plane with diameter $d > 0$. Among the disks that contain F there is a unique closed disk D of least diameter. The diameter of D is at most $2d/\sqrt{3}$.*

For a simple proof, see J. von Sz. Nagy [**6**]. A proof for the case in which F is finite may be found in Rademacher and Toeplitz [**5**, p. 103]. For a proof based on Helly's theorem, see [**3**, p. 14]. For historical remarks and other references, see [**1**].

The question concerning the possible values of c can now be answered as follows.

THEOREM 3. *Let $F_1, \ldots, F_m$ be disjoint nonempty finite sets and let U be an open subset of the plane that contains a convex neighborhood of each of the sets F_j. If and only if $c > 2/\sqrt{3}$ do there always exist disjoint polygonal arcs $A_1, \ldots, A_m$ (possibly degenerate) such that $F_j \subset A_j \subset U$ and* $\operatorname{diam} A_j \leqq c \operatorname{diam} F_j$ *$(j = 1, \ldots, m)$.*

Proof. Let F denote the union of the singleton sets F_j, and let the others (if any) be numbered $F_1, \ldots, F_k$. By Theorem 2, for any $c > 2/\sqrt{3}$ there exist open disks D_j such that $F_j \subset D_j$ and $\operatorname{diam} D_j < c \operatorname{diam} F_j$ $(j = 1, \ldots, k)$. The intersection of D_j with any convex neighborhood of F_j is convex, hence F_j is contained in a single component of $D_j \cap U - F$. The conclusion then follows from Theorem 1. Example 1 shows that the conclusion need not hold when $c \leqq 2/\sqrt{3}$.

COROLLARY. *Under the hypotheses of Theorem 3, let d be an upper bound to the diameters of the sets F_j. In order that there exist disjoint regions $R_1, \ldots, R_m$, each of diameter less than or equal to d', such that $F_j \subset R_j \subset U$ $(j = 1, \ldots, m)$, it is sufficient, and in general necessary, that d' be greater than $2d/\sqrt{3}$.*

In §3 we shall show that the problem of separating members of a very large class of compact zero-dimensional sets is reducible to the case of finite sets.

If each of the sets F_j has at most two elements, arcs (and regions) of smaller diameter can always be found. Indeed, Theorem 1 implies immediately that in this case c may be assigned any value greater than 1. In particular, disjoint pairs of points within distance ε of each other can always be separated by disjoint regions of diameter less than ε. (An easy special case of this result was used in the construction described in Chapter 18 of [**4**].) However, in contrast to Theorem 3, it is remarkable that for 2-element

sets the lower bound $c = 1$ is actually attained. One can even demand that the diameter of each arc be attained only at its endpoints.

THEOREM 4. *Let $\{p_j, q_j\}$ $(j = 1, \ldots, m)$ be disjoint pairs of points in the plane and let U be an open set that contains circular arcs, each less than a semicircle, joining p_j to q_j. Then there exist disjoint polygonal arcs $A_1, \ldots, A_m$ such that A_j joins p_j to q_j, $A_j \subset U$, and $|p - q| < |p_j - q_j|$ for all p, q in A_j with $\{p, q\} \neq \{p_j, q_j\}$.*

This result cannot be deduced from Theorem 1, since disjoint arcs certainly cannot be constructed within the spanning disks in case two of the pairs span the same disk. However, disjoint arcs can be constructed within regions that differ only slightly from these disks and have the same diameter, as will be shown in §4.

Theorems 3, 4, and Example 1 leave open the question of what values of c are permissible in case all but one of the sets has at most two points.

2. Proof of Theorem 1. The proof is by induction on m, but it calls for repeated modification of arcs previously drawn. The constructions required are entirely elementary; each step consists in joining points by an arc within a given region, or in extending an arc to a simple closed curve in such a way as to enclose given points in its interior. The order of the steps at each stage is governed by a certain combinatorial tree.

Until §4, the unmodified term "arc" will always mean a polygonal arc regarded as a point set; that is, an image of $[0, 1]$ by a piece-wise linear homeomorphism.

The theorem is true when $m = 1$, so we may assume that $m > 1$, that the theorem is true when the number of sets is $m - 1$, that each of the sets F_j is nonempty and contained in a single component of $D_j \cap U$, and that the sets have been numbered so that the diameter of D_m is maximal. By hypothesis, there exist disjoint arcs $A_1, \ldots, A_{m-1}$ such that

$$F_j \subset A_j \subset D_j \cap U - F_m \qquad (j = 1, \ldots, m - 1) \tag{1}$$

and

$$A_1, \ldots, A_{m-1} \text{ are disjoint.} \tag{2}$$

Let C denote the boundary of D_m. Since $A_j \cap C$ is finite, each component of the set $G = D_m - \bigcup_{j=1}^{m-1} A_j$ is bounded by a finite number of arcs of the following four types:

(i) arcs whose endpoints belong to C and whose intermediate points belong to D_m and to some A_j,

(ii) arcs having one endpoint on C and whose other points belong to D_m and to some A_j and include an endpoint of A_j,

(iii) arcs A_j that are entirely contained in D_m, and

(iv) circular subarcs of C.

Let $\alpha_1, \ldots, \alpha_N$ denote the arcs of type (i), if any. No two of these intersect in D_m and each of them separates two adjacent components of G. By induction on $N \geqq 0$, the number of components of G is seen to be exactly $N + 1$. Since each α_i is contained in some A_j, the diameter of α_i is less than that of D_j, and therefore of D_m. Let C_i denote the shorter of the two circular arcs into which C is divided by the endpoints of α_i. For $1 \leqq i \leqq N$ let R_i denote the (unique) component of G that lies in the interior of the simple closed curve $\alpha_i \cup C_i$ and whose boundary ∂R_i contains α_i. Denote the one remaining component of G by R_0. Let us say that R_j *lies immediately above* R_i if $R_j \neq R_i$ and $\alpha_i \subset \partial R_j$. Let $R_0, \ldots, R_N$ denote also the vertices of a graph T in which two vertices are joined by an edge if and only if one of the corresponding regions lies immediately above the other. For each $i > 0$ there is a unique j such that R_j lies immediately above R_i. Moreover, R_i can be joined to R_0 by a chain of regions each of which (after R_i) lies immediately above its predecessor. It follows that T is a tree; in fact, since R_0 is a distinguished region, a rooted tree. (It is easy to see that an arbitrary rooted tree can

be realized in this manner.) Assign to each vertex of T, and to the corresponding region, a *rank* equal to the number of edges needed to join it to R_0. Let us say that R_i is *occupied* if $F_m \cap R_i \neq \varnothing$. Let T_0 denote the smallest connected subgraph of T that contains all of the occupied vertices. The number K of edges of T_0 is then a well defined nonnegative integer associated with any set of arcs $A_1, \ldots, A_{m-1}$ that satisfy (1) and (2). Let us call K the *index* of $A_1, \ldots, A_{m-1}$. The proof of Theorem 1 will evidently be complete when we have established the following two lemmas.

LEMMA 1. *Given any set of arcs $A_1, \ldots, A_{m-1}$ that satisfy* (1) *and* (2) *and have index $K > 0$, it is possible to modify one of them so as to obtain a new set of arcs that satisfy* (1) *and* (2) *and have index $K - 1$.*

LEMMA 2. *Given any set of arcs $A_1, \ldots, A_{m-1}$ that satisfy* (1) *and* (2) *and have index $K = 0$, there exists an arc A_m such that $F_m \subset A_m \subset U \cap G$.*

Proof of Lemma 1. Let R_i be an occupied region of maximum rank. Then $i > 0$ and there is a point $p \in R_i \cap F_m$ and a point $q \in R_k \cap F_m$ for some $k \neq i$. q must lie outside the simple closed curve $\alpha_i \cup C_i$ (otherwise R_k would be an occupied region of higher rank than R_i), whereas p lies inside it. By hypothesis, p can be joined to q by an arc $A \subset D_m \cap U$, and A must meet α_i. If A directed from p to q meets some $\alpha_j \subset \partial R_i$ before it meets α_i, then R_i must lie immediately above R_j and A must reenter R_i before it can meet α_i. Since $\alpha_j \subset U$ we can modify A so that it stays in $U \cap D_m$ and does not meet α_j at all. Similarly, if A meets an arc of type (ii) or (iii) before it meets α_i, then A can be modified within $U \cap D_m$ so as to avoid this arc by going around one of its endpoints or by paralleling it from a point near its first to one near its last intersection with this arc. We conclude that p can be joined to a point r of α_i by an arc in $(R_i \cap U) \cup r$. Because $\alpha_i \subset U$ it follows that every point of $F_m \cap R_i$ belongs to the same component of $R_i \cap U$. Call this component H.

Let A_j be the arc that contains α_i and let β be a subarc of α_i that contains no point of F_j. Because $\alpha_i \subset \partial H$ we can join the two points of $\partial\beta$ (the endpoints of β) by an arc $\gamma \subset H \cup \partial\beta$ constructed in such a way that the simple closed curve $\beta \cup \gamma$ encloses all the points of $F_m \cap R_i$ in its interior. Since $\alpha_i \subset D_j$ and $\operatorname{diam} D_j \leqq \operatorname{diam} C$, it follows that C_i and the region bounded by $\alpha_i \cup C_i$ are both contained in D_j (see Remark below). Therefore $\gamma - \partial\beta \subset D_j \cap U \cap G - F_m$. Thus if we replace β by γ, A_j will be modified in such a way that (1) and (2) are maintained. Moreover, $C_1, \ldots, C_N$ will remain unchanged, and likewise the graph T, although two of its vertices now represent different regions than they did before. The vertex R_i is no longer occupied, but the adjacent vertex in T_0 is occupied, whether or not it was before. Thus the effect of replacing A_j by $(A_j - \beta) \cup \gamma$ is to leave T unchanged but to delete from T_0 the vertex R_i and the adjacent edge. The modified set of arcs therefore has index $K - 1$.

Proof of Lemma 2. If $K = 0$, then F_m is entirely contained in one of the regions R_i, $0 \leqq i \leqq N$. Any two points of F_m can be joined by an arc in $D_m \cap U$, and this arc can be chosen so that it does not meet any arc of type (i), (ii), or (iii). Thus F_m is contained in a single component of $R_i \cap U$. Hence there exists an arc A_m such that

$$F_m \subset A_m \subset D_m \cap U - \bigcup_{j=1}^{m-1} A_j.$$

REMARK. The assumption that D_j is circular and not larger than D_m was used at the point indicated (and in the definition of C_i). Similar reasoning applies if each D_j is a square open interval not larger than D_m. (To insure that $A_j \cap \partial D_m$ is always finite one may conveniently restrict attention to arcs made up of segments that are neither vertical nor horizontal.) However, Example 2 shows that the neighborhoods D_j cannot be taken to be arbitrary rectangular intervals.

3. Application to a lemma of Goffman. In this section we shall show that Theorem 1 provides just what is needed to complete the proof, in the case $n = 2$, of a certain lemma of C. Goffman [**2**, Lemma 4] as modified by H. E. White, Jr. [**7**].

A plane set is said to be *sectionally zero dimensional* if its complement contains a dense set of horizontal lines and a dense set of vertical lines. A (2-dimensional) *p-set* is a compact 2-cell whose boundary is the union of a finite number of horizontal or vertical line segments. In the following, a rectangle will always mean a 2-dimensional closed interval.

The 2-dimensional case of the Goffman–White lemma is contained in the following more precise theorem, which also generalizes the corollary to Theorem 3 stated above. (Example 2 shows that the proof of the Goffman–White lemma given in [2] does not work in the case $n = 2$, at least for any obvious interpretation of the sets Q_j introduced there. The bound $2d/\sqrt{3}$ in the following theorem is sharp; it improves the value $\sqrt{2}d$ given by Goffman.)

THEOREM 5. *Let $K_1, \ldots, K_m$ be disjoint sectionally zero dimensional closed sets, each of diameter less than d, and let U be an open subset of E^2 that contains the convex hull of each of the sets K_j. There exist disjoint p-sets $P_1, \ldots, P_m$ such that $K_j \subset \operatorname{int} P_j$, $P_j \subset U$, and* $\operatorname{diam} P_j < 2d/\sqrt{3}$ *for $j = 1, \ldots, m$.*

Proof. By Theorem 2, each of the sets K_j (which we may assume nonempty) is contained in an open disk D_j of diameter less than $2d/\sqrt{3}$. Choose $\delta > 0$ such that the δ-neighborhood D'_j of each D_j also has diameter less than $2d/\sqrt{3}$ and such that any rectangle of diameter less than δ that meets one of the sets K_j meets no other and is contained in $D_j \cap U$. Divide the plane into rectangles of diameter less than δ by means of horizontal and vertical lines that do not meet any of the sets K_j. Let $P'_{j1}, \ldots, P'_{jm_j}$ be those rectangles of this subdivision that meet K_j, and let $P_{j1}, \ldots, P_{jm_j}$ be slightly smaller concentric rectangles whose interiors still cover K_j. Let V denote the complement relative to U of the union of all of the (disjoint) rectangles P_{ji}. Let F_j be a set formed by choosing just one point from each of the (nonempty) sets $P_{ji} \cap K_j$ ($i = 1, \ldots, m_j$). Since $F_j \subset K_j \subset D_j \cap U$ and $D_j \cap U$ contains the convex hull of K_j, F_j is contained in a single component of $D_j \cap U$. By Theorem 1 there exist disjoint arcs $A_1, \ldots, A_m$ such that

$$F_j \subset A_j \subset D_j \cap U \qquad (j = 1, \ldots, m).$$

In case $m_j = 1$ replace A_j by the singleton F_j, and when $m_j > 1$ let A_j be chosen so that both of its endpoints belong to F_j and that each of its segments that meets F_j is horizontal. Choose rectangles Q_{ji} contained in and similar to P_{ji} each centered at the (unique) element of $F_j \cap P_{ji}$ and small enough so that the sets

$$B_j = A_j \cup \bigcup_{i=1}^{m_j} Q_{ji} \qquad (j = 1, \ldots, m)$$

are disjoint and that, when $m_j > 1$, Q_{ji} meets A_j in a horizontal line segment. Let Φ be a piece-wise linear automorphism of E^2 that maps each Q_{ji} onto P_{ji} affinely and is equal to the identity outside and on the boundaries of all of the rectangles P'_{ji}. Then the sets

$$B'_j = \Phi(B_j) = \Phi(A_j) \cup \bigcup_{i=1}^{m_j} P_{ji}$$

are disjoint. Because the norm of Φ is less than δ, we have $\Phi(A_j) \subset D'_j$, hence $B'_j \subset D'_j \cap U$. When $m_j = 1$ let $A''_j = F_j$, and when $m_j > 1$ let A''_j be an arc composed of horizontal or vertical line segments that differs from $\Phi(A_j)$ only in V and approximates $\Phi(A_j)$ so closely that the sets $B''_j = A''_j \cup \bigcup_{i=1}^{m_j} P_{ji}$ are disjoint and $B''_j \subset D'_j \cap U$ ($j = 1, \ldots, m$). When $m_j = 1$ take $P_j = P_{j1}$, and when $m_j > 1$ take P_j to be the union of the rectangles $P_{j1}, \ldots, P_{jm_j}$ and the closed rectangular ε-neighborhoods of A''_j. Then if ε is chosen sufficiently small, $P_1, \ldots, P_m$ will be disjoint p-sets with $P_j \subset D'_j \cap U$, and we will have $K_j \subset \operatorname{int} P_j$, $P_j \subset U$, and $\operatorname{diam} P_j < 2d/\sqrt{3}$ for $j = 1, \ldots, m$.

4. Proof of Theorem 4. In this section we shall use the term "arc" to designate any homeomorphic image of $[0, 1]$ that is composed of a finite number of straight line segments *or circular arcs.*

Let $F_j = \{p_j, q_j\}$ ($j = 1, \ldots, m$) be disjoint 2-element subsets of the plane. Let D_j be the interior of

the circle C_j spanned by F_j, let $d_j = \operatorname{diam} D_j$, and put $F = \bigcup_{j=1}^m F_j$. Let $N_\delta(p)$ denote the open disk with radius δ and center p, and let $\bar{N}_\delta(p)$ denote its closure.

If $0 < \varepsilon < d_j/2$, it is clear that for any sufficiently small $\delta > 0$ the modified disk $[D_j \cup N_\delta(p_j)] - \bar{N}_\varepsilon(q_j)$ will have the same diameter as D_j. It is an easy exercise in trigonometry to verify that this is true if and only if $\delta \leqq d_j - \sqrt{d_j^2 - \varepsilon^2}$, and therefore that it is sufficient to take $\delta < \varepsilon^2/2d_j$. Hence if

(i) $0 < \varepsilon < \frac{1}{2}|p - q|$ for all p, q in F with $p \neq q$, and

(ii) $0 < \delta < \varepsilon^2/2 \max d_j$,

and if $I_j = \{i: 1 \leqq i \leqq m,\ i \neq j,\ C_i = C_j\}$, then each of the sets

$$R_j = [D_j \cup \bigcup_{i \in I_j} N_\delta(p_i)] - \bigcup_{i \in I_j} \bar{N}_\varepsilon(q_i) \qquad (j = 1, \ldots, m), \tag{3}$$

is an open region of diameter d_j. The boundary ∂R_j of R_j is a simple closed curve (composed of a finite number of circular arcs) through the points p_j and q_j. Only at diametrically opposite points of C_j on the boundary of R_j is the diameter of ∂R_j attained. If in addition

(iii) ε is less than the distance from any point $p \in F$ to any circle C_j for which $p \notin C_j$,

then the following lemma holds.

LEMMA 3. *If $j < m$, $d_j \leqq d_m$, $\{p, q\} \subset \partial R_m \cap (R_j \cup F_j)$, and $p \neq q$, then just one of the two arcs into which p and q divide ∂R_m has diameter less than d_m, and this arc β is contained in $R_j \cup \partial R_j$. If $C_j \neq C_m$, then $\beta \subset R_j \cup F_j$.*

Proof. Suppose first that $C_j \neq C_m$. There are only four cases in which the set $E = \partial R_m \cap (R_j \cup F_j)$ has more than one point:

CASE (a): $C_k = C_j$ for some $k \neq j$ and C_j is either internally or externally tangent to C_m at p_k. Then $E = C_m \cap N_\delta(p_k)$.

CASE (b): $C_k = C_m$ for some $k \neq m$ and C_j is internally tangent to C_m at q_k. Then $E = \partial N_\varepsilon(q_k) \cap D_j$.

CASE (c): $C_k = C_m$ for some $k \neq m$ and C_j is externally tangent to C_m at p_k. Then $E = \partial N_\delta(p_k) \cap D_j$.

CASE (d): C_j meets C_m in two points. In this case ∂R_j and ∂R_m meet in just two points and E is the open arc of ∂R_m that lies in R_j between them, plus either endpoint that may happen to belong to F_j.

In each of these four cases E is a connected subset of ∂R_m. In no case does E contain a pair of opposite points of C_m, whereas $\partial R_m - E$ always contains such a pair. In the first three cases E is contained in R_j; in all cases $E \subset R_j \cup F_j$. One of the two arcs into which p and q divide ∂R_m is contained in E and has diameter less than d_m; the other one has diameter d_m.

Suppose finally that $C_j = C_m$. Then E is the union of two disjoint open circular arcs: $C_m \cap N_\delta(p_m)$ and $\partial N_\varepsilon(q_j) \cap D_j$. One of the two arcs into which p and q divide ∂R_m does not contain q_m. This arc is contained in $R_j \cup \partial R_j$. It has diameter less than d_m because it contains no pair of opposite points of C_m. The other arc has diameter d_m.

Now let U be an open subset of the plane such that $F \subset U$ and assume that, for each j, p_j and q_j belong to the boundary of one component of $D_j \cap U$. Then it is clear that any sufficiently small $\varepsilon > 0$ will satisfy not only (i) and (iii) but also

(iv) for each j there is an arc A_j such that

$$F_j \subset A_j \subset (D_j \cap U - \bar{N}_\varepsilon(F - F_j)) \cup F_j.$$

We shall prove the following theorem inductively.

(T_m). *If $F_j = \{p_j, q_j\}$ $(j = 1, \ldots, m)$ are disjoint 2-element subsets of an open set U in the plane, if each F_j is contained in the boundary of one component of $D_j \cap U$, and if ε and δ satisfy* (i) (ii) (iii) *and* (iv),

then there exist disjoint polygonal arcs $A_1, \ldots, A_m$ *such that*

$$F_j \subset A_j \subset (R_j \cap U) \cup F_j \qquad \text{for } j = 1, \ldots, m$$

where R_j *is defined by* (3).

Proof. First renumber the sets so that $d_j \leqq d_m$ for all $j < m$. By hypothesis there exists an arc A_1 such that $F_1 \subset A_1 \subset (D_1 \cap U) \cup F_1$. This establishes (T_1), since $R_1 = D_1$ in case $m = 1$. Assume now that $m > 1$, that (T_{m-1}) is true, and that $F_1, \ldots, F_m$, U, ε, δ satisfy the hypotheses of (T_m). Then $F_1, \ldots, F_{m-1}$, $U - \bar{N}_\varepsilon(F_m)$, ε, δ satisfy the hypotheses of (T_{m-1}), and so there exist disjoint polygonal arcs $A_1, \ldots, A_{m-1}$ such that $F_j \subset A_j \subset (R'_j \cap U') \cup F_j$ $(j = 1, \ldots, m-1)$, where $U' = U - \bar{N}_\varepsilon(F_m)$ and R'_j is the region defined by (3) with i restricted to values less than m. By (iv) there exists a polygonal arc A_m such that $F_m \subset A_m \subset (D_m \cap U - \bar{N}_\varepsilon(F - F_m)) \cup F_m$, and we may assume that no segment of A_m is parallel to any of the segments that compose $A_1, \ldots, A_{m-1}$. Since $R'_j \cap U' \subset R_j \cap U$ for $j < m$, and $D_m \cap U - \bar{N}_\varepsilon(F - F_m) \subset R_m \cap U$, it follows that

$$F_j \subset A_j \subset (R_j \cap U) \cup F_j \quad (j = 1, \ldots, m), \tag{4}$$

$$A_1, \ldots, A_{m-1} \text{ are disjoint, and} \tag{5}$$

$$\text{the set } M = A_m \cap \bigcup_{j=1}^{m-1} A_j \text{ is finite.} \tag{6}$$

We shall prove (T_m) by showing that if $A_1, \ldots, A_m$ are polygonal arcs that satisfy (4) (5) and (6), and if $M \neq \varnothing$, then it is always possible to modify one of the arcs in such a way that (4) and (5) still hold and M is reduced to a proper subset.

Regard each of the arcs A_j as directed from p_j to q_j. If A_i contains an endpoint r of some other arc A_j, or if A_i and A_j intersect at a point $r \notin F_i \cup F_j$ but do not cross there, then the intersection at r can be eliminated by modifying A_i within an arbitrarily small neighborhood of r. Let us call such an intersection (and modification) trivial. We shall consider two further types of modification.

Suppose that A_i and A_j have only nontrivial intersections, that $r \in A_i \cap A_j$, and that the subarc α of A_j joining r to one of the endpoints s of A_j is contained in R_i and contains no point of the set $A = \bigcup_{k \neq i,j} A_k$. Let α' denote the subarc of α joining s to the point r' of A_i nearest to s along α. Then for any sufficiently small $\delta' > 0$ the closed rectangular δ'-neighborhood V of α' will be contained in $R_i \cap U - A$ and its boundary will be a simple closed polygon that meets A_j in just one point and meets A_i in just two points p and q. Choose such a number δ' and let A' be that one of the two arcs into which p and q divide ∂V that does not meet A_j. Substitute A' for the part of A_i between p and q. After this modification of A_i, (4) and (5) will still be satisfied and M will be reduced to $M - \{r'\}$. Let us call such a modification of A_i one of type I.

Suppose next that $j < m$, that A_j and A_m have only nontrivial intersections, that α is a subarc of A_j, that A_m crosses α t times, $t \geqq 2$, and that the first and last of these crossings are in opposite directions relative to A_j. Suppose further that the shortest subarc α' of α that contains all of these crossings is contained in R_m. Let A denote the union of the arcs A_k with $k \neq j, m$. Then for any sufficiently small $\delta' > 0$ the closed rectangular δ'-neighborhood V of α' will be contained in $R_m \cap U - A$, and its boundary will be a simple closed polygon. Moreover, $V \cap A_j$ will consist of a single arc and $V \cap A_m$ of t disjoint arcs, all having their endpoints on ∂V. Choose such a number δ' and let p and q be the first and last points of A_m on ∂V. Then one of the two arcs into which p and q divide ∂V does not meet A_j. Substitute this arc for the part of A_m between p and q. Then (4) and (5) will still be satisfied and M will be reduced to a subset having at least two fewer elements. Let us call such a modification of A_m one of type II.

Now let $A_1, \ldots, A_m$ satisfy (4), (5), and (6), and assume that $M \neq \varnothing$. Let r and r' be the first and last points of A_m that belong to M, say $r \in A_j$ and $r' \in A_{j'}$. We shall show that by a trivial modification of A_m, A_j, or $A_{j'}$, or by one of type I or II, we can always reduce M to a proper subset.

Assume that none of the intersections of A_m with A_j or $A_{j'}$ is trivial, and that none of the six inclusions $p_m r \subset R_j$, $p_j r \subset R_m$, $rq_j \subset R_m$, $p_{j'}r' \subset R_m$, $r'q_{j'} \subset R_m$, or $r'q_m \subset R_{j'}$ holds (otherwise a modification of type I of one of these arcs is possible). Then there is a last point p of $p_j r$ on ∂R_m, and a first point q of rq_j on ∂R_m. Let α denote the subarc pq of A_j. Similarly, let α' denote the subarc of $A_{j'}$ that joins the last point p' of $p_{j'}r'$ on ∂R_m to the first point q' of $r'q_{j'}$ on ∂R_m. Let s be the last point of A_m on α, and let s' be the first point of A_m on α'. Assume from now on that A_m crosses α in the same direction at r and s, and that it crosses α' in the same direction at s' and r' (otherwise a modification of A_m of type II is possible). We shall show that these hypotheses are untenable.

By Lemma 3 one of the two arcs β into which p and q divide ∂R_m has diameter less than d_m and is contained in $R_j \cup \partial R_j$. Hence the simple closed curve $\alpha \cup \beta$ is contained in $R_j \cup \partial R_j$, as well as in $R_m \cup \partial R_m$, and its interior R is contained in $R_j \cap R_m$. Suppose A_m crosses $\alpha \cup \beta$ in the outward direction at r. Then all points of A_m between p_m and r belong to R, and p_m must belong to β. Hence $p_m \in R_j$ (by definition in case $C_j = C_m$ and by Lemma 3 in case $C_j \neq C_m$), and so the entire arc $p_m r$ is contained in R_j, contrary to hypothesis. Therefore the crossings at r and s must both be inwards relative to $\alpha \cup \beta$. It follows similarly that q_m must belong to β and that all points of A_m between s and q_m belong to R. We must have $C_j \neq C_m$ (otherwise q_m would lie outside ∂R_j), and therefore $q_m \in R_j$ (by Lemma 3). Hence $sq_m \subset R_j$. We must have $j \neq j'$ (otherwise $r'q_m \subset sq_m \subset R_{j'}$, contrary to hypothesis), and therefore $r' \in R$. Consequently p' and q' belong to $\beta - \{p, q\}$ and $\alpha' - \{p', q'\} \subset R$. Let β' denote the subarc of β joining p' to q'. Then $\alpha' \cup \beta'$ is a simple closed curve contained in $R \cup \partial R$; its interior R' is contained in R. Since β' has diameter less than d_m it follows from Lemma 3 that $\beta' \subset R_{j'} \cup \partial R_{j'}$. Therefore $\alpha' \cup \beta' \subset R_{j'} \cup \partial R_{j'}$ and $R' \subset R_{j'}$. Because s lies outside of $\alpha' \cup \beta'$, the crossings at s' and at r' must both be inward relative to $\alpha' \cup \beta'$. Therefore all points of A_m between r' and q_m belong to R', and q_m must belong to β'. We must have $C_{j'} \neq C_m$ (otherwise q_m would lie outside $\partial R_{j'}$), and therefore $q_m \in R_{j'}$ (by Lemma 3). Hence the entire arc $r'q_m$ is contained in $R_{j'}$, contrary to hypothesis. This final contradiction establishes (T_m). Theorem 4 is an immediate corollary.

References

1. L. M. Blumenthal and G. E. Wahlin, On the spherical surface of smallest radius enclosing a bounded subset of n-dimensional euclidean space, Bull. Amer. Math. Soc., 47 (1941) 771–777.
2. C. Goffman, One-one measurable transformations, Acta Math., 89 (1953) 261-278.
3. H. Hadwiger, H. Debrunner, and V. L. Klee, Combinatorial Geometry of the Plane, Holt, Rinehart and Winston, New York, 1964.
4. J. C. Oxtoby, Measure and Category, Springer, New York–Heidelberg–Berlin, 1971.
5. H. Rademacher and O. Toeplitz, The Enjoyment of Mathematics, Princeton University Press, Princeton, 1957.
6. J. von Sz. Nagy, Über einen Satz von H. Jung, Jber. Deutsch. Math.-Verein., 24 (1915) 390–392.
7. H. E. White, Jr., The approximation of one-one measurable transformations by measure preserving homeomorphisms, Proc. Amer. Math. Soc., 44 (1974) 391–394.

Department of Mathematics, Bryn Mawr College, Bryn Mawr, PA 19010.

THE PRODIGAL INTEGRAL

GEORGE R. SELL

I. Introduction. One of the well-known theorems of integral calculus is the following:

Let $f(t)$ and $g(t)$ be real or complex-valued functions with $f'(t) = g(t)$ for all t in the real line R. If $g(t)$ is of class $C^n(R)$, then $f(t)$ is of class $C^{n+1}(R)$.

In other words, integration increases the order of differentiation.

In this note we shall present a construction which may seem to be a paradox. Specifically, we shall construct two continuous complex-valued functions $F(x_1, x_2)$ and $G(x_1, x_2)$ with the following properties:

(i) F and G are doubly periodic in x_1 and x_2, i.e., both F and G satisfy

$$F(x_1, x_2) = F(x_1 + 2\pi, x_2) = F(x_1, x_2 + 2\pi) \tag{1}$$

for all $x = (x_1, x_2) \in R^2$. This means that F and G are, in fact, complex-valued functions defined on the torus $T^2 = S^1 \times S^1$, and x_1 and x_2 are the respective angular variables.

(ii) G is of class $C^4(T^2)$ but F is not.

(iii) If $f(t) = F(v_1 t, v_2 t)$ and $g(t) = G(v_1 t, v_2 t)$, then for appropriately chosen $v = (v_1, v_2)$ one has $f'(t) = g(t)$ for all $t \in R$.

(iv) The ratio v_1/v_2 is irrational so that the values of F and G on T^2 are uniquely determined by $f(t)$ and $g(t)$.

By saying that G is of class $C^k(T^2)$ we mean, as usual, that G and all derivatives up to at least order k are continuous on T^2. The curiosity of this construction arises from conditions (ii) and (iii). Even though $f(t)$ — as a function of t — has more derivatives than $g(t)$, the extensions F and G do not share this property.

The construction we present in Section 2 is based on the theory of Fourier series for almost periodic functions. We begin with $v = (v_1, v_2) \in R^2$ given where v_1/v_2 is irrational, and we study almost periodic functions $h(t)$, which have a Fourier series expansion

$$h(t) = \sum_n h_n e^{in \cdot vt},$$

where $n = (n_1, n_2)$ is a vector with integral entries and $n \cdot v = n_1 v_1 + n_2 v_2$. This function $h(t)$ can be considered as a function defined for $t \in R$ or as a function with domain $\gamma(v) = \{(v_1 t, v_2 t) \in T^2 : t \in R\}$. In the latter case the domain of h is a dense subset of the torus, and under certain conditions h admits an extension to a function $H(x_1, x_2)$ on T^2 of class $C^k(T^2)$. The conditions we shall study will guarantee that the function $h(t)$ is uniformly continuous on $\gamma(v)$. Since $\gamma(v)$ is dense in T^2, the extension of h to a function H on T^2 is uniquely determined by asking that H be continuous. We shall also want to know when this extension H is of class $C^k(T^2)$.

The problem of showing that G is of class $C^4(T^2)$, and F is not, is based on a number theoretic property of badly approximable numbers. The argument showing the existence of a pair of numbers $v = (v_1, v_2)$ with the properties we require is presented in Section 3. We would like to express our appreciation to Frank Meyer and J. Ian Richards for their kind and helpful assistance on the number theoretic aspects of this paper.

II. The construction. We begin with a brief review of the theory of Fourier series. The space $L_2(T^2)$ is the collection of all measurable complex-valued functions $F(x) = F(x_1, x_2)$ with $\int_{T^2} |F|^2 d\mu < \infty$, where μ is the surface area measure on T^2. This L_2-space is, of course, the same as the L_2-space on the square $[0, 2\pi] \times [0, 2\pi]$ with the usual Lebesgue measure. Every function $F \in L_2(T^2)$ has a Fourier series

$$F(x) = \sum_n f_n e^{in \cdot x}, \tag{2}$$

where $\Sigma_n |f_n|^2 < \infty$.

Let $v = (v_1, v_2) \in R^2$ be fixed where v_1/v_2 is irrational. For each $F \in L_2(T^2)$ define $f(t)$ by

$$f(t) = F(v_1 t, v_2 t). \tag{3}$$

Under appropriate conditions it happens that $f(t)$ is an almost periodic function and $f(t)$ has a Fourier

series

$$f(t) = \sum_n f_n e^{in\cdot vt}, \tag{4}$$

where the coefficients f_n are determined by F and equation (2). Let us now examine this relationship in greater detail. We define H_v to be Hilbert space of all almost periodic functions $f(t)$ which have a Fourier series expansion $f(t) = \Sigma_n f_n e^{in\cdot vt}$ with $\Sigma_n |f_n|^2 < \infty$. The mapping $\Phi: F \to f$ that maps $\Sigma_n f_n e^{in\cdot x}$ onto $\Sigma_n f_n e^{in\cdot vt}$ then defines a unitary equivalence between $L_2(T^2)$ and H_v.

Thus the unitary equivalence Φ between $L_2(T^2)$ and H_v is defined in terms of the Fourier series. We would like to know under what conditions the mapping Φ can be realized by equation (3). That is, given that $f = \Phi(F)$, where is it true that $f(t) = F(v_1 t, v_2 t)$ for all t? This question is very delicate and it is connected with the question of pointwise convergence of a Fourier series. Recall that an L_2-function is prescribed only up to a set of measure zero. Furthermore, the set on which equation (3) is to hold, i.e., the set

$$\gamma(v) = \{(v_1 t, v_2 t) \in T^2 : t \in R\},$$

is a set of measure zero in T^2. It is not important for our purposes to discuss this question in this generality. Instead, we shall restrict ourselves to continuous functions.

The realization of Φ by means of equation (3) has two aspects. First, given a continuous doubly periodic function $F(x_1, x_2)$, when does equation (3) hold for all $t \in R$ where $f = \Phi(F)$? Secondly, given a continuous function $f(t)$ in H_v, when does equation (3) hold for all $t \in R$ where $F = \Phi^{-1}(f)$? A sufficient condition for the realization of Φ by means of equation (3) is given in terms of the Fourier coefficients.

PROPOSITION 1. *Assume that $\Sigma_n |f_n| < \infty$. Then the functions F and f, defined by equations* (2) *and* (4), *respectively, are continuous, and equation* (3) *is valid for all $t \in R$.*

The proof is not difficult. The function $e^{in\cdot x}$ is doubly periodic and continuous on T^2. Since $\Sigma_n |f_n| < \infty$, the series $\Sigma_n f_n e^{in\cdot x}$ converges uniformly and therefore the sum $F(x)$ is continuous. Likewise, $f(t) = \Sigma_n f_n e^{in\cdot vt}$ is continuous, and finally, it is clear that equation (3) holds for all $t \in R$, since the Fourier series in (2) and (4) converge uniformly. Q.E.D.

The question of the differentiability of F can be answered in a similar fashion. First, consider the formal derivatives of F, i.e.,

$$\frac{\partial F}{\partial x_1} = \sum_n n_1 f_n e^{in\cdot x}, \quad \frac{\partial F}{\partial x_2} = \sum_n n_2 f_n e^{in\cdot x}.$$

It follows from Proposition 1 that if $\Sigma_n |n_i f_n| < \infty$, then $\partial F / \partial x_i$ is continuous on T^2, $i = 1, 2$. Similar considerations apply for higher order derivatives. Let $|n|$ denote $\max(|n_1|, |n_2|)$. We then have the following result.

PROPOSITION 2. *Assume that $\Sigma_n |n|^k |f_n| < \infty$ for some integer $k = 1, 2, \ldots$. Then the functions F and f, defined by equations* (2) *and* (4), *respectively, are of class C^k, and equation* (3) *is valid for all $t \in R$.* (We mean here, of course, that F is of class $C^k(T^2)$ and f is of class $C^k(R)$.)

For each integer $k = 0, 1, 2, \ldots$, we let D_k denote the collection of all functions F which have a Fourier series given by equation (2) and such that

$$\sup_n |n|^k |f_n| < \infty.$$

We then have the following

PROPOSITION 3. *$D_2 \subseteq L_2(T^2)$ and $D_{k+3} \subseteq C^k(T^2) \subseteq D_k$ for $k = 0, 1, 2, \ldots$.*

In order to prove this we first observe that the series $\Sigma_{n\neq 0}|n|^{-K}$ is convergent if and only if the double integral $\iint_\Delta r^{-K}dxdy$ is convergent where $r=(x^2+y^2)^{1/2}$ and $\Delta=\{(x,y)\in R^2: x^2+y^2\geqq 1\}$. The last integral can easily be evaluated by use of polar coordinates, and we conclude that $\Sigma_{n\neq 0}|n|^{-K}$ is convergent whenever $K>2$. In particular, the series $\Sigma_{n\neq 0}|n|^{-3}$ is convergent.

Now if $F\in D_{k+3}$, then there is a $B<\infty$ such that

$$|n|^k|f_n|\leqq B|n|^{-3}$$

for $n\neq 0$. It follows then from Proposition 2 that $F\in C^k(T^2)$. Likewise, if $F\in D_2$ then

$$|f_n|^2\leqq(B|n|^{-2})^2=B^2|n|^{-4}$$

for $n\neq 0$, and therefore $\Sigma_n|f_n|^2<\infty$, or $F\in L_2(T^2)$.

Let us now show that $C^0(T^2)\subseteq D_0$. If $F\in C^0(T^2)$ then $F\in L_2(T^2)$ and $\Sigma_n|f_n|^2<\infty$. Hence, $\sup_n|f_n|^2<\infty$, or $\sup_n|f_n|<\infty$, i.e., $F\in D_0$. Likewise, if $F\in C^1(T^2)$, then the derivatives $\partial F/\partial x_1$ and $\partial F/\partial x_2$ are in $L_2(T^2)$ and therefore $\Sigma_n|n_1f_n|^2<\infty$ and $\Sigma_n|n_2f_n|^2<\infty$. Hence, $\sup_n|n_1f_n|<\infty$ and $\sup_n|n_2f_n|<\infty$. Since

$$\begin{aligned}\sup_n|n|\,|f_n|&=\sup_n\,[\max(|n_1|,|n_2|)\,|f_n|]\\&=\max[\sup_n|n_1f_n|,\ \sup_n|n_2f_n|],\end{aligned}$$

we see that $F\in D_1$. By repeating this argument one concludes that $C^k(T^2)\subseteq D_k$. Q.E.D.

We are now prepared to study the differential equation $f'(t)=g(t)$ in more detail. Let us first look at this equation formally and later we shall study the convergence problems. If f and g have the Fourier series $f(t)=\Sigma_nf_ne^{in\cdot vt}$ and $g(t)=\Sigma_ng_ne^{in\cdot vt}$, then $f'(t)=\Sigma_n in\cdot vf_ne^{in\cdot vt}$. Consequently, $f'(t)=g(t)$ implies that $g_0=0$ and

$$f_n=-i\frac{g_n}{n\cdot v},\qquad n\neq 0. \tag{5}$$

We see then that the properties of v can become very crucial. Not only does v_1/v_2 need to be irrational (which prevents $n\cdot v$ from vanishing for $n\neq 0$) but also one must be concerned that $n\cdot v$ does not become too small. This is the well-known small divisors problem. Thus for each integer $K=1,2,\ldots,$ we define A_K to be the collection of $v=(v_1,v_2)\in R^2$ such that there is a $d=d(v)>0$ with the property that

$$|n\cdot v|\geqq d|n|^{-K}$$

for all $n\neq 0$. The connection between the sets A_K and the differential equation $f'(t)=g(t)$ is described in the next result.

PROPOSITION 4. *Let $v\in A_K$ for some integer $K\geqq 1$ and let $G\in D_{K+k}$ for some integer $k\geqq 3$. Then there is an $F\in D_k$ such that*

$$f'(t)=g(t)\qquad\textit{for all}\qquad t\in R,$$

where $f=\Phi(F)$ and $g=\Phi(G)$.

The proof is based on equation (5). Since $v\in A_K$ one has

$$|n|^k\,|f_n|=\frac{|n|^k|g_n|}{|n\cdot v|}\leqq\frac{|n|^k|g_n|}{d|n|^{-K}}=\frac{1}{d}|n|^{K+k}|g_n|$$

for $n\neq 0$. The proof now follows from Propositions 1, 2, and 3. Q.E.D.

This suggests that the function f, or better, the corresponding doubly periodic function F can have fewer derivatives than G. We shall now construct such an example. For this construction we shall use a $v \in A_4$ with the property that $v \notin A_3$. The existence of such a number v will be discussed in the next section.

Let G be defined by $g_0 = 0$ and $g_n = |n|^{-7}$ for $n \neq 0$, and let F be defined by $f_0 = 0$ and

$$f_n = -i\frac{g_n}{n \cdot v}$$

for $n \neq 0$. Let $g = \Phi(G)$ and $f = \Phi(F)$.

THEOREM. *The following statements are valid*:

(A) $G \in D_7 \subsetneq C^4(T^2)$ *and* $g(t) = G(v_1t, v_2t)$ *for all* $t \in R$.

(B) *If* $v \in A_4$, *then* $F \in D_3 \subsetneq C^0(T^2)$. *Also one has* $f(t) = F(v_1t, v_2t)$ *and* $f'(t) = g(t)$ *for all* $t \in R$.

(C) *If* $v \in A_4$ *and* $v \notin A_3$, *then* $f \notin D_4$. *In particular, one has* $F \notin C^4(T^2)$.

The proof that $G \in D_7$ is a straight-forward computation. The remaining assertions in statement (A) follow from Propositions 2 and 3.

The proof of statement (B) follows directly from Propositions 1, 3, and 4.

The proof of statement (C) deserves some comment. If $v \notin A_3$ then for every integer $k \geqq 1$ there is an n_k such that

$$|n_k \cdot v| \leqq k^{-1}|n_k|^{-3}.$$

Therefore for $n = n_k$ one has

$$|f_n| = \frac{|g_n|}{|n \cdot v|} \geqq \frac{|n|^{-7}}{k^{-1}|n|^{-3}} = k|n|^{-4},$$

which implies that $F \notin D_4$. By Proposition 3, one then has $F \notin C^4(T^2)$. Q.E.D.

III. A number theoretic result. The sets A_K defined in the last section are related to the collection of badly approximable numbers, which in turn is the complement of the collection of Liouville numbers. For each integer $K \geqq 1$ we let B_K denote the collection of all real numbers α such that there is a $c = c(\alpha) > 0$ with the property that

$$\left|\alpha - \frac{p}{q}\right| \geqq \frac{c}{q^K}$$

for all rational numbers p/q, where p and q are relatively prime. The sets B_K are increasing in K, and it is easily seen that one has $(1, \alpha) \in A_K$ if and only if $\alpha \in B_{K+1}$, for $K = 1, 2, \ldots$. We shall now show that the sets B_K are strictly increasing in K for $K \geqq 2$. If we choose then an $\alpha \in B_5$ with $\alpha \notin B_4$, it follows that $v = (1, \alpha) \in A_4$ and $v \notin A_3$.

Our argument will be based on the theory of continued fractions. We will present here the essential properties of this theory which we shall need for our proof. The reader is referred to [2] for more details.

For $n = 0, 1, 2, \ldots$ let a_n denote an integer and assume that $a_n \geqq 1$ for $n = 1, 2, \ldots$. Define the quantity $[a_0, a_1, \ldots, a_n]$ inductively by setting

$$[a_0, a_1, \ldots, a_n] = p_n/q_n$$

where $p_{-1} = 1$, $q_{-1} = 0$, $p_0 = a_0$, $q_0 = 1$, $p_1 = p_{-1} + a_1p_0$, $q_1 = q_{-1} + a_1q_0$, or in general,

$$p_{n+1} = p_{n-1} + a_{n+1}p_n, \quad q_{n+1} = q_{n-1} + a_{n+1}q_n, \tag{6}$$

for $n = 0, 1, 2, \ldots$. (For example, $[0, 1, 2, 3] = 7/10$.) Recall that every irrational number α can be

represented as a continued fraction

$$\alpha = [a_0, a_1, a_2, \ldots] = \lim_{n\to\infty} [a_0, a_1, \ldots, a_n],$$

where the a_n are integers with $a_n \geqq 1$ for $n = 1, 2, \ldots$. Conversely, given any collection of integers a_n, $n = 0, 1, 2, \ldots$, with $a_n \geqq 1$ for $n = 1, 2, \ldots$, then the $\lim_{n\to\infty} [a_0, a_1, \ldots, a_n]$ exists and is an irrational number α.

The two properties we shall need are the following:

Proposition 5. *If* $\alpha = [a_0, a_1, a_2, \ldots]$ *and*

$$\left| \alpha - \frac{p}{q} \right| < \frac{1}{2q^2}$$

for some rational number p/q, *then* $p/q = p_n/q_n$ *for some* $n = 0, 1, 2, \ldots$.

Proposition 6. *For* $n = 0, 1, 2, \ldots$ *one has*

$$\frac{1}{q_n^2(a_{n+1}+2)} \leqq \left| \alpha - \frac{p_n}{q_n} \right| \leqq \frac{1}{q_n^2 a_{n+1}}. \tag{7}$$

Let us now show that the sets B_K are strictly increasing in K for integral values $K \geqq 2$. Fix $K \geqq 2$ and define $\alpha = [a_0, a_1, a_2, \ldots]$ by $a_0 = 0$, $a_1 = 3$ and let a_{n+1} for $n \geqq 1$ be defined inductively by

$$a_{n+1} = q_n^{K-2} - 2. \tag{8}$$

The number α is well-defined. One computes (p_1, q_1) by (6). Next one computes a_2 by (8). (Note that $q_1 = 3$ so $a_2 \geqq 1$.) Then one goes back to (6) to compute (p_2, q_2) where $q_2 > q_1$. The process then continues *ad infinitum.*

Let us now show that $\alpha \in B_K$ and $\alpha \notin B_{K-1}$. It follows from (7) that

$$\frac{1}{q_n^K} \leqq \left| \alpha - \frac{p_n}{q_n} \right|$$

for $n = 0, 1, 2, \ldots$. If $p/q \neq p_n/q_n$ for some $n = 0, 1, 2, \ldots$, then it follows from Proposition 5 that

$$\left| \alpha - \frac{p}{q} \right| \geqq \frac{1}{2q^2} \geqq \frac{1}{2q^K}.$$

Hence $\alpha \in B_K$. On the other hand, inequality (7) implies that

$$\left| \alpha - \frac{p_n}{q_n} \right| \leqq \frac{1}{q_n^2(q_n^{K-2}-2)} \leqq \frac{2}{q_n^K}$$

for $n = 1, 2, \ldots$. The last term, of course, becomes smaller than any term of the form δ/q_n^{K-1} for any preascribed $\delta > 0$. Hence $\alpha \notin B_{K-1}$.

Research for this paper was supported in part by NSF Grant No. GP-38955.

References

1. W. Maak, Fastperiodische Funktionen, Springer, Berlin, 1950.
2. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, Clarendon Press, Oxford, 1962.
3. E. Hewitt and K. A. Ross, Abstract Harmonic Analysis, Vol. I, II, Springer, Berlin, 1963, 1970.

School of Mathematics, University of Minnesota, Minneapolis, MN 55455.

IMMERSIONS AND MOD-2 QUADRATIC FORMS

LOUIS H. KAUFFMAN AND THOMAS F. BANCHOFF

1. Introduction. We are going to consider an easily visualizable classification problem in topology and its close relationship with a corresponding classification problem in algebra. It is standard practice in topology to begin with the geometry and then mirror part of its structure in algebra. Usually it requires care to find a mirror which captures just enough geometry so that the corresponding algebra problems are both accessible and relevant.

It occasionally happens that this process illuminates both the algebra and the geometry. This is the case with our topic. Hence we obtain a particularly nice way to learn about two things simultaneously — immersions of surfaces and mod-2 quadratic forms.

It is difficult to talk about two things at once. We shall therefore begin with the topology, show how it leads to quadratic forms, and then discuss quadratic forms in more detail. By doing this, we obtain geometric proofs and interpretations for the basic algebraic identities which underlie the theory of mod-2 quadratic forms.

This paper is relatively self-contained except for a few facts about surfaces that are summarized in sections 3 and 4 and some facts about homology in section 6. In such cases, we have tried to make the facts geometrically plausible.

A word about quadratic forms: Consider a polynomial function $f(x, y) = ax^2 + bxy + cy^2$ with a, b, and c real numbers. The locus $f(x, y) =$ constant represents a conic section in the plane and it is a standard exercise to determine the geometric form of this conic by changing variables to eliminate the xy term in the expression. A modern approach is to consider the plane as the vector space R^2 of ordered pairs $v = (x, y)$ of real numbers, and to think of $f: R^2 \to R$ as a function on R^2. We then can define a bilinear form on R^2 by setting $\langle v, w\rangle = f(v+w) - f(v) - f(w)$. Conversely, given a bilinear form on R^2, we may obtain a quadratic form $q: R^2 \to R$ by setting $q(v) = \frac{1}{2}\langle v, v\rangle$. Similarly in any field where $1+1 \neq 0$, there is a one-to-one correspondence between bilinear forms and quadratic forms. But our geometric study will lead us to quadratic forms on a vector space V over the field Z_2 with two elements, where $1+1=0$. Although for any such form $q: V \to Z_2$ we get a bilinear form $\langle v, w\rangle = q(v+w) + q(v) + q(w)$, the quadratic form is no longer completely determined by the bilinear form. Thus the study of mod-2 quadratic forms involves a particular subtlety. We shall see corresponding phenomena mirrored in the study of immersions, and this correspondence is the main point of this paper.

The paper is organized as follows. Section 2 discusses immersions of circles into R^2 and the two sphere, S^2, and shows how consideration of the two sphere leads to mod-2 phenomena. Sections 3 and 4 discuss surfaces and their immersions. Letting $\mathscr{C}(M)$ denote the set of embedded curves on a surface M, we obtain, for each immersion $f: M \to S^2$, a function $N(f): \mathscr{C}(M) \to Z_2$. This function measures how curves on M are immersed into S^2. Section 4 introduces an invariant, $B(f)$, for immersions of surfaces by examining $N(f)$ on the boundary curves. We study immersions of punctured disks up to an equivalence relation called image homotopy. This section introduces some basic homotopies (handle sliding and permutation) which will be used later. In section 6 we show how $N(f): \mathscr{C}(M) \to Z_2$ leads to a quadratic form $q(f): \mathscr{H}(M) \to Z_2$ where $\mathscr{H}(M)$ is the mod-2 homology group of M. In section 7 we discuss certain homotopies of immersions and show how they correspond to isomorphisms of quadratic forms. It is then easy to explain the classification of mod-2 forms. Section 8 completes the classification of immersions of surfaces up to image homotopy.

A concise exposition of mod-2 quadratic forms may be found in [1, pp. 52–56]. For more information about topology and quadratic forms, the reader may enjoy looking at [**3**] and [**5**]. We remark that our results are related to the research article by Rourke and Sullivan [**6**].

2. Regular closed curves in the plane and the sphere. In order to approach the study of surfaces we

have to understand closed curves. A closed curve may be thought of as the path traced out by a point moving continuously in the plane R^2, so that the point ends where it began. We may describe such a path α by writing $\alpha(\theta)$ to indicate the position of the point that corresponds to the angle θ on the circle S^1, so that $\alpha(0)=\alpha(2\pi)$. We will restrict ourselves to *regular curves* $\alpha: S^1 \to R^2$. These curves are also called *immersions* of the circle. They are defined by the condition that the velocity vector $\alpha'(\theta)$ varies continuously and is non-zero for all values of θ and $\alpha'(0)=\alpha'(2\pi)$. As a point traces out a regular curve, there is a well-defined continuously turning direction vector $\alpha'(\theta)/\|\alpha'(\theta)\|$ for each θ, and as θ goes from 0 to 2π, this unit vector goes a certain number of times around the unit circle in a counterclockwise direction. We define this integer to be the *degree* $D(\alpha)$ of the regular closed curve. Figure 1 indicates the degrees of a number of regular closed curves with arrows indicating which way each curve is to be traversed.

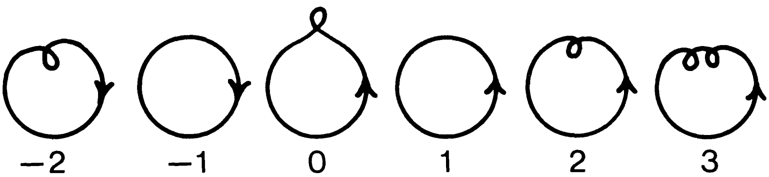


FIG. 1

The degree of a curve does not change if we deform the curve slightly so that the tangent directions move in a continuous manner. We call a one-parameter family $\alpha_t: S^1 \to R^2$, $0 \leqq t \leqq 1$, of regular curves a *regular homotopy* if the tangent vectors $\alpha_t'(\theta)$ to these curves vary continuously as θ and t change, and we say that α and $\bar{\alpha}$ are *regularly homotopic* if there is such a deformation beginning with $\alpha=\alpha_0$ and ending with $\bar{\alpha}=\alpha_1$. In this case we write $\alpha \simeq \bar{\alpha}$. The definition of regular homotopy is set up so that if α and $\bar{\alpha}$ are regularly homotopic, then $D(\alpha)=D(\bar{\alpha})$.

Our story really begins with a theorem of Hassler Whitney ([7], p. 279) which establishes the converse of this last statement. His main result shows that the integer $D(\alpha)$ completely characterizes the curves that are regularly homotopic to α.

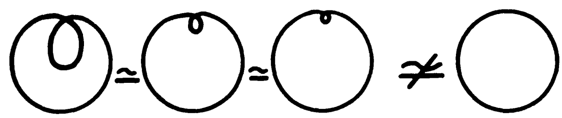


FIG. 2

THEOREM 2.1 (Whitney–Graustein). *If α and $\bar{\alpha}$ are two regular curves in R^2 with $D(\alpha)=D(\bar{\alpha})$ then α and $\bar{\alpha}$ are regularly homotopic.*

The condition $\alpha'(\theta) \neq 0$ for all θ indicates that for any θ_0, the curve is approximated by the tangent line, at least for a small interval about the value θ_0, so that as θ moves through this interval the regular curve α is one-to-one, with no double points. It is possible however for the whole regular curve to have double points. A double point is a point in R^2 that is the image of two distinct points on the circle. That is, the point p satisfies the condition: $p=\alpha(\theta_0)=\alpha(\theta_1)$ for some $0 \leqq \theta_0 \leqq \theta_1 < 2\pi$ and $\alpha^{-1}(p)=\{\theta_0, \theta_1\}$. Similarly, we may speak of points whose pre-image consists of a finite number of points on the circle (triple points, etc.). We call a double point a *normal crossing* if the tangent lines to the curve at the two points are different. We say that a regular curve is *normal* if all of its self-intersections are double points with normal crossings. Any regular curve may be deformed by a regular homotopy into a normal curve. During a regular homotopy, the number of normal crossings may change, but the parity will remain the same. (See Figure 3 for an example of a regular homotopy

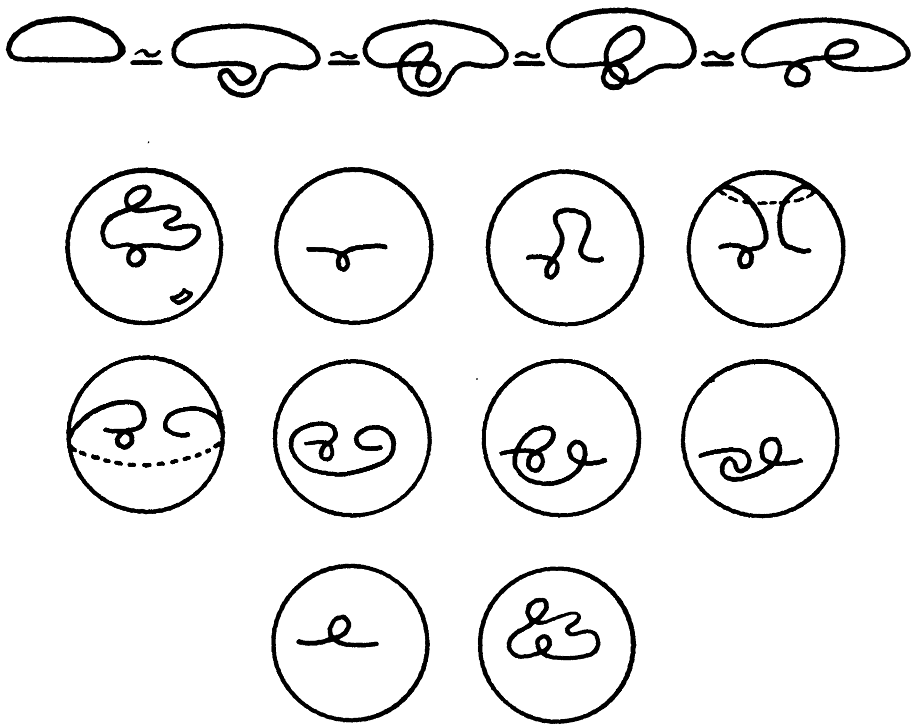


FIG. 3

which introduces two new normal crossings.) Thus if α and $\bar{\alpha}$ are normal curves with $\alpha \simeq \bar{\alpha}$, then the number of normal crossings of α and the number of normal crossings of $\bar{\alpha}$ are either both even or both odd. (Regularity fails at the last instant in the parity-changing deformation illustrated in Figure 2.) If α is normal, we define the *crossing number* $N(\alpha)$ of α to be the number of normal crossings reduced modulo two.

Whitney [7] established a relation between the degree of a normal curve and the number of normal crossings:

THEOREM 2.2. *If α is a normal curve on R^2, then $D(\alpha)$ and the number of normal crossings of α have opposite parity. That is, $D(\alpha)+1$ reduced modulo 2 equals $N(\alpha)$.*

We now wish to consider closed curves not in the plane but in S^2, the unit sphere in R^3. The notions of regular curve, regular homotopy, and normal curve carry over to the case where we map to S^2 rather than R^2, but now there is an additional kind of deformation available, in which we swing a loop over the back of the sphere. This is illustrated in Figure 3. The first part of the figure shows how to create or destroy two normal crossings. Note that when this occurs in the plane the two tiny loops so created contribute oppositely to the degree since the tangent vector turns in opposite directions on the two loops. In the plane it is not possible to change a $+1$ loop to a -1 loop by a regular homotopy. However, this can be done in S^2. The second part of Figure 3 shows how to switch such a loop on a curve without disturbing the rest of the curve. For this figure we ask the reader to imagine that he or she is looking down towards the surface of a transparent sphere. As the curve on the sphere is deformed, part of it swings over the back of the sphere (dotted lines). In the intermediate stages of the regular homotopy we have not drawn the entire curve. What is not drawn remains stationary. A deformation such as this obliterates the difference between curves 3 and 1 of Figure 1, or between curves 2 and 0.

Although the notion of degree as we have defined it does not apply to a normal curve on the sphere, we can still speak about the normal crossing number, $N(\alpha)$. The number $N(\alpha)$ is an invariant of the regular homotopy class of α. Whitney's result shows that the crossing number completely classifies normal curves on the sphere up to regular homotopy:

THEOREM 2.3. *For normal curves α and $\bar{\alpha}$ on S^2, $\alpha \simeq \bar{\alpha}$ if and only if $N(\alpha)=N(\bar{\alpha})$.*

3. Immersions of the annulus in the sphere. In this paper we want to use the properties of regular curves to study immersions of surfaces with boundary into S^2. Perhaps the simplest such surface from our point of view is the *annulus*, which we may think of as an interval of concentric circles in the plane, with radii varying from r_0 to r_1. By an *immersion* of the annulus into R^2 or S^2, we mean a mapping $\alpha(\theta, r)$, $0 \leqq \theta \leqq 2\pi$, $r_0 \leqq r \leqq r_1$, such that the partial derivative vectors $\partial\alpha/\partial\theta$, $\partial\alpha/\partial r$ are non-collinear vectors at each point which vary continuously as r and θ change. This definition guarantees that the annulus is mapped in a locally one-to-one way. Note that it follows that, for each fixed r, the curve $\alpha(\theta, r)$ is an immersion of the circle. For example, a small strip neighborhood of a regular curve in the plane is the image of an immersion of the annulus. If we would deform this center curve by a regular homotopy $\alpha_t(\theta, r)$, then we could obtain a similar one-parameter family $\alpha_t(\theta, r)$ of immersions of the entire annulus. We say that such a family of immersions is a *regular homotopy* of the annulus if for every r, the family $\alpha_t(\theta, r)$ is a regular homotopy of curves and if the partial derivative vectors $\partial\alpha_t/\partial\theta$ and $\partial\alpha_t/\partial r$ vary continuously as θ, r and t change.

As in the case of curves, we are interested in classifying regular homotopy classes of immersions of the annulus into the sphere, and we may use our results on curves to give a complete classification.

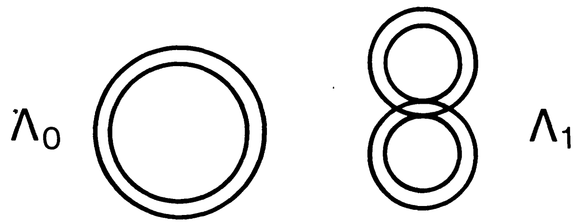


FIG. 4

THEOREM 3.1. *Any immersion of the annulus in S^2 is regularly homotopic to an immersion whose image is a strip neighborhood of either a circle or of a figure eight.* (See Figure 4.)

Proof. Any immersion of the annulus may be shrunk down to a strip neighborhood of the center curve. (This involves the tubular neighborhood theorem of differential topology. See Milnor, *Topology from the Differentiable Viewpoint,* The University Press of Virginia, (1965) p. 46.) We may then use the results of the previous section to find a regular homotopy of this curve to a circle or a figure eight carrying along the strip neighborhood of the curve.

There is one additional subtlety that makes the theory of immersions of the annulus different from that of curves and that has to do with *orientation.* The partial derivative vectors $\partial\alpha/\partial\theta$ and $\partial\alpha/\partial r$ are tangent vectors to the sphere that are assumed to be non-collinear, so their cross-product either points out of the ball or into the ball at every point of the annulus. In the first case we say that α preserves orientation and in the second we say that α reverses orientation. The property of preserving or reversing orientation does not change during a regular homotopy, so our classification theorem may be stated more precisely as follows:

THEOREM 3.2. *Two immersions of the annulus into the sphere are regularly homotopic if and only if their center curves are regularly homotopic and both immersions either preserve or reverse orientation.*

Henceforth we shall assume that all of our immersions of surfaces are orientation preserving.

4. Immersions of surfaces with boundary into the sphere. An *orientable surface with boundary* is obtained by removing a finite number of non-overlapping discs from an orientable surface such as a sphere, a torus, or more generally, a sphere with handles. An annulus, for example, may be described as a sphere with two discs removed. The disc itself is a sphere with a disc removed. A basic theorem in the study of orientable surfaces states that any such surface may be obtained by starting with a polygonal region in the plane and identifying certain pairs of boundary edges (see for example [4], Chapter 1). A torus may be described in this way as a square with opposite sides identified, and if we remove a quarter disc about each corner of the square before identifying the opposite sides, we obtain a torus with a disc removed, which we shall refer to as a *punctured torus* (see Figure 5).

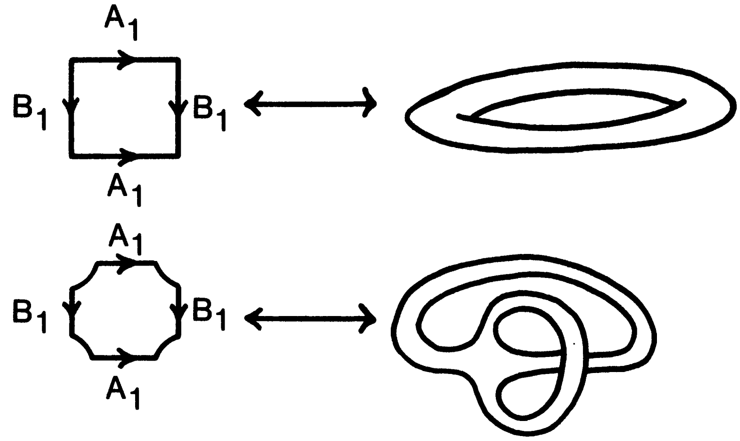


FIG. 5

Although it is not possible to find a locally one-to-one mapping of the entire torus into the plane, it is possible to produce several different immersions of the punctured torus into the plane. To do this we *attach bands* to the square without corners by finding immersions of half of an annulus $\alpha(\theta, r)$, $0 \leqq \theta \leqq \pi$, $r_0 \leqq r \leqq r_1$, so that the ends match up with the sides that are to be identified. In a region about any point of the punctured torus we may then find parameters so that the partial derivative vectors are non-collinear and so that they vary continuously as the parameters change. We call such a mapping an *immersion* of the punctured torus into the plane (or into S^2).

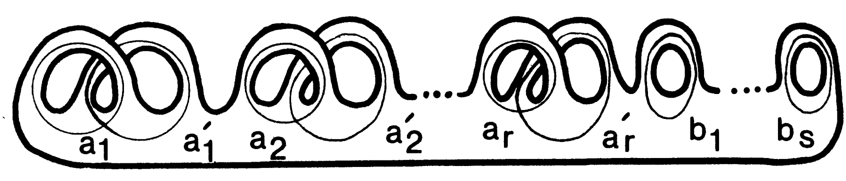


FIG. 6

We obtain the *center curve of a band* by taking the center curve of the half annulus and connecting its endpoints to the center of the square by a one-to-one arc to form an immersion of the circle. These remarks apply equally well to an arbitrary surface. In the general case the procedure described above leads to a surface with boundary curves as illustrated in Figure 6. In Figure 6 the collection of center curves to the bands is given by the set

$$\{a_1, a'_1, a_2, a'_2, \ldots, a_r, a'_r, b_1, \ldots, b_s\}.$$

A band is said to be *untwisted* if its center curve is regularly homotopic to an embedded curve (that

is, a curve with no self-crossings) and to be *twisted* if it is regularly homotopic to a normal curve with just one crossing. In Figure 7 we indicate four immersions of the punctured torus: T_{00} with both bands untwisted, T_{01} and T_{10} with one band twisted and one band untwisted, and T_{11} with both bands twisted.

As in the case of the annulus, we define a *regular homotopy* of immersions of a surface M with boundary to be a one-parameter family of immersions $f_t : M \to S^2$ so that for some region about any point we may find a parametrization so that the partial derivative vectors change continuously as t varies.

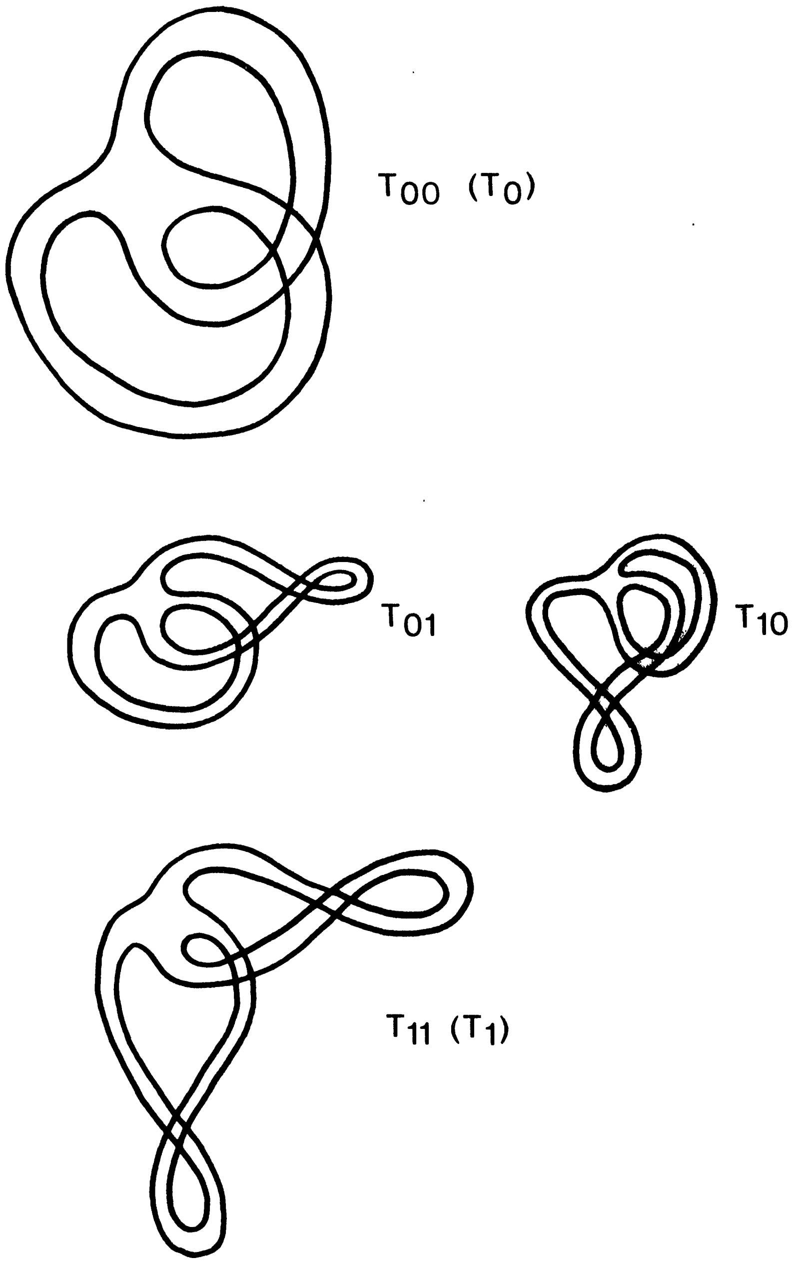


FIG. 7

The regular homotopy classification of surface immersions is not hard. Our next theorem is a generalization of Theorems 3.1 and 3.2. Since the proof is similar to the proofs of 3.1 and 3.2, we shall omit it.

THEOREM 4.1. *Let $f: M \to S^2$ be an orientation preserving immersion of a surface M with band curves $\{\alpha_1, \ldots, \alpha_k\}$. Another immersion, $g: M \to S^2$, is regularly homotopic to f if and only if $N(f \circ \alpha_i) = N(g \circ \alpha_i)$ for each $i = 1, 2, \ldots, k$.*

Thus an immersion is determined up to regular homotopy by the crossing numbers of its band curves. There are four regular homotopy classes of immersions of the punctured torus. An example of an immersion in each class is given in Figure 7. The images in the figure are denoted T_{00}, T_{01}, T_{10}, and T_{11}.

The immersions T_{01} and T_{10} differ very little in appearance. In fact, the distinction we are making between them depends upon more than their images. We are assuming that maps $f: M \to S^2$ and $g: M \to S^2$ are given so that $f(M) = T_{01}$, $g(M) = T_{10}$. Furthermore, if α and β are the band curves on M then we also assume that $N(f \circ \alpha) = 0$, $N(f \circ \beta) = 1$ while $N(g \circ \alpha) = 1$, $N(g \circ \beta) = 0$. By changing a map without changing its image one can produce non-regularly homotopic immersions with the same image. For example, let $h: M \to M$ be a homeomorphism (that is, a one-to-one, onto, continuous mapping with continuous inverse) of the punctured torus that switches the two bands. If h satisfies the same differentiability criteria that we imposed upon an immersion, then so will the composition $f \circ h$. Hence $f \circ h: M \to S^2$ is also an immersion. We assumed that h interchanged the bands and therefore $N((f \circ h) \circ \alpha) = 1$ and $N((f \circ h) \circ \beta) = 0$. Thus, while $T_{01} = f(M) = f \circ h(M)$, we see that f and $f \circ h$ are not regularly homotopic. In fact, by Theorem 4.1, $f \circ h$ is regularly homotopic to g. Thus, depending upon the maps representing them, the images T_{01} and T_{10} may represent distinct or equivalent immersions.

We wish to concentrate on the images of immersions. For this purpose it is useful to say that an orientation preserving homeomorphism $h: M \to M$ (M any orientable surface) is a *diffeomorphism* if h satisfies the local derivative conditions for an immersion. It then follows that if $f: M \to S^2$ is an immersion and $h: M \to M$ is a diffeomorphism, then $f \circ h: M \to S^2$ is also an immersion. Note that f and $f \circ h$ have identical images.

We are going to study an equivalence relation on immersions called *image homotopy*. Intuitively, two immersions are image homotopic if there is a regular homotopy between their images. For example, consider the bottom line of Figure 8; it illustrates an image homotopy between T_{00} and T_{01}. Thus image homotopy is weaker than regular homotopy. Since our rigorous definition of image homotopy is a bit technical, we defer it to the end of this section. The reader may wish to look forward into the rest of the paper before examining the exact concept of image homotopy.

In setting up algebraic invariants of image homotopy it will be useful to consider curves embedded in a given surface. Let $\mathscr{C}(M)$ denote the collection of curves $\alpha: S^1 \to M$ with α an embedding. Given an immersion $f: M \to S^2$ and $\alpha \in \mathscr{C}(M)$ we obtain a regular curve $f \circ \alpha: S^1 \to S^2$. Thus we may define

$$N(f): \mathscr{C}(M) \to Z_2 \quad \text{by} \quad N(f)(\alpha) = N(f \circ \alpha).$$

In the next section $N(f)$ will be applied to the boundary curves of M. We then turn to arbitrary curves and see how this leads to a connection between image homotopy and mod-2 quadratic forms.

Notation. Since T_{00}, T_{01}, and T_{10} are image homotopic, we shall use the notation T_0 for T_{00} and T_1 for T_{11} in all later sections.

Here is the promised definition:

DEFINITION 4.2. Two immersions f, $g: M \to S^2$ are image homotopic ($f \approx g$) if there is a diffeomorphism $h: M \to M$ so that $f \circ h$ is regularly homotopic to g.

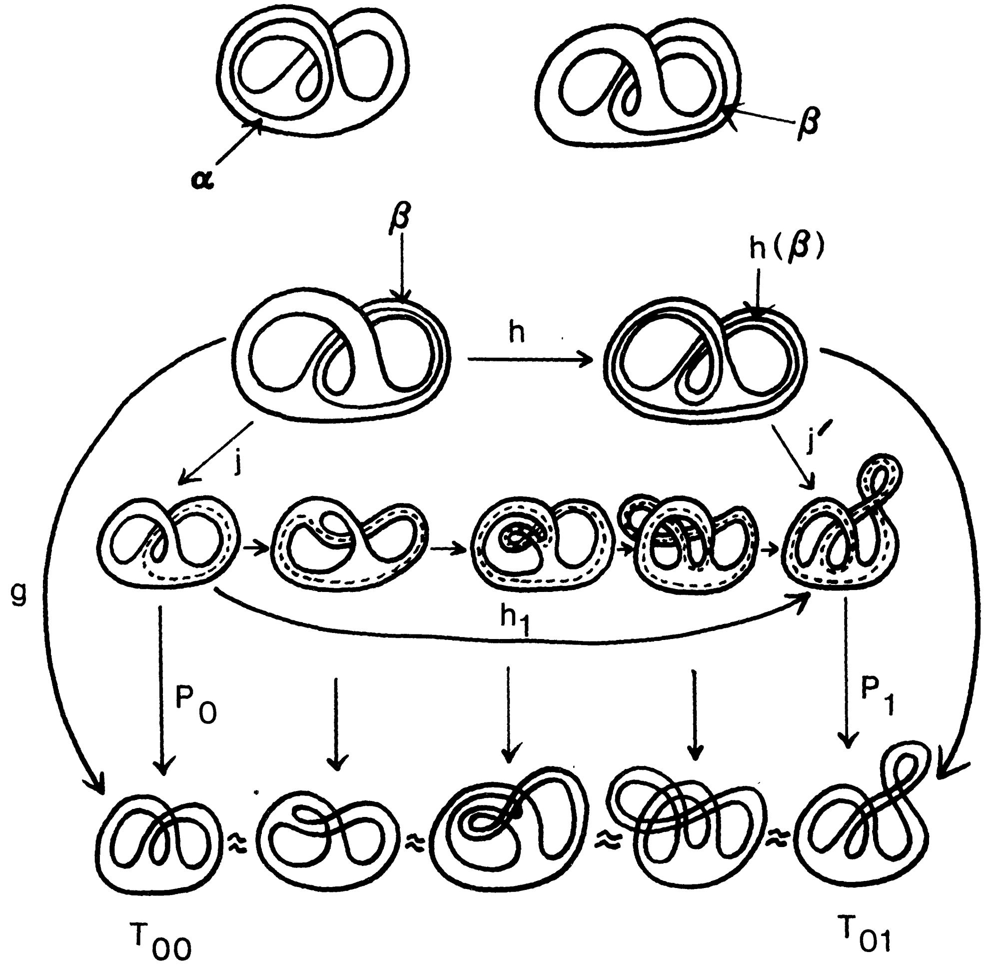


FIG. 8

While this definition may seem hard to visualize, in practice this is not the case. A diffeomorphism of a surface may often be viewed as the result of a deformation. View Figure 8. It illustrates (middle row) an embedding of the punctured torus in R^3 and a deformation through embeddings to a new punctured torus in R^3. More precisely, we are given a time-parameter family of diffeomorphisms $h_t : R^3 \to R^3$ so that $h_0 = \text{identity}$, and an embedding $j : M \to R^3$. Then $j(M)$ represents the embedded torus at time $t = 0$ and $h_1(j(M))$ represents the torus at time $t = 1$. There is an obvious map $j' : M \to R^3$ so that $j'(M) = h_1(j(M))$ and so that j' embeds the disk and maps each band to the corresponding embedded band. Define $h : M \to M$ by the formula $h(x) = (j'^{-1} \circ h_1 \circ j)(x)$. As the picture suggests, $h(\alpha) = \alpha$ while $h(\beta)$ is a sort of combination of α and β, where α and β are the band curves on M. The bottom part of the figure illustrates a sequence of projections $p_t : h_t(j(M)) \to R^2$ (or S^2). In each case, $p_t \circ h_t \circ j(M)$ is an immersed surface. Let $f = p_1 \circ j'$ and $g = p_0 \circ j$. Then f and g are immersions with $N(g)(\alpha) = N(g)(\beta) = 0$ while $N(f)(\alpha) = 0$, $N(f)(\beta) = 1$. However, by our construction we see that $N(f \circ h)(\alpha) = 0 = N(f \circ h)(\beta)$. Hence, by Theorem 4.1, $f \circ h$ is regularly homotopic to g.

We put this in a nutshell by saying that T_{00} and T_{01} are image homotopic. Intuitively, the image homotopy is pictured in the bottom line of Figure 8. Any such picture may be unfolded as we have

done, to produce a diffeomorphism h. We shall use this convention from now on, drawing image homotopies as in the bottom of Figure 8 and leaving the unfolding to the reader.

5. The boundary invariant. Recall that our surfaces have boundary. Each boundary component of a surface M may be viewed as a curve on M. If C is a boundary component of M, then we may choose an embedding $c: S^1 \to M$ so that $c(S^1) = C$. In this way we regard each boundary curve as an element of $C(M)$. Given an immersion $f: M \to S^2$, we may compute the crossing number, $N(f)(C) = N(f \circ c)$, for each boundary component C. For a given component C, this number does not depend upon the choice of embedding $c: S^1 \to M$.

DEFINITION 5.1. The boundary invariant, $B(f) \in Z$, of an immersion $f: M \to S^2$ is the total number of boundary curves $C \subset M$ such that $N(f)(C) = 1$.

LEMMA 5.2. *If f and g are image homotopic immersions of a surface M into S^2, then $B(f) = B(g)$.*

Proof. First suppose $f = g \circ h$ where $h: M \to M$ is a diffeomorphism. Since h only permutes the boundary components, $B(f) = B(g)$. If f is regularly homotopic to g then $f|_C$ is regularly homotopic to $g|_C$ for each boundary curve C. Hence $N(f)(C) = N(g)(C)$ and therefore $B(f) = B(g)$. Since it is sufficient to check these two cases, we conclude that if f is image homotopic to g, then $B(f) = B(g)$.

LEMMA 5.3. *For any surface immersion $f: M \to S^2$, the boundary invariant is even. In particular, if M has only one boundary component, then $B(f) = 0$.*

Proof. Given an immersion $f: M \to S^2$ we may assume that each boundary curve is immersed with normal crossings, and that the immersions of distinct boundary curves cross each other normally. Let the total crossing number, $T(f) \in Z$, be the total number of crossings (self and mutual) among the immersed boundary curves. Since any two curves in S^2 intersect in an even number of points, it is clear that $B(f) \equiv T(f)$ (modulo 2).

We shall prove that $T(f) \equiv 0$ (modulo 2) by induction on the number of bands necessary to represent M. If there are no bands, then M is a disk and so there is a regular homotopy of the boundary curve to a small circle about a point. Such a circle has crossing number zero, and so the original crossing number was even. For a surface with one or more bands, we may eliminate a band by cutting across it as in Figure 9. The total crossing number is unaffected by this operation. Hence the lemma follows by induction.

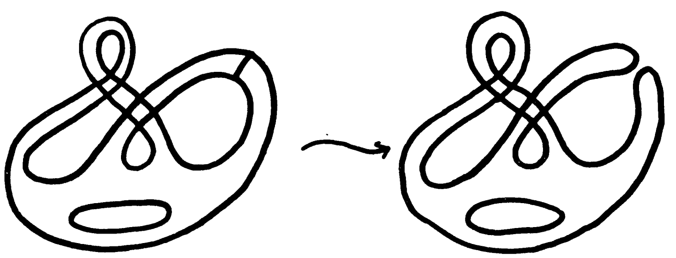


FIG. 9

The boundary invariant easily distinguishes certain immersions. For example, let X and Y denote the immersions in Figure 10. Then $B(X) = 2$ while $B(Y) = 4$. Hence X is not image homotopic to Y.

In fact, immersions of punctured disks are classified by the boundary invariant:

PROPOSITION 5.4. *Let M be a punctured disk with k holes, represented as a disk with k attached bands. Two immersions f, $g: M \to S^2$ are image homotopic if and only if $B(f) = B(g)$.*

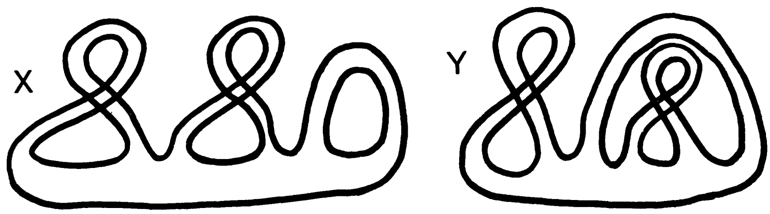


FIG. 10

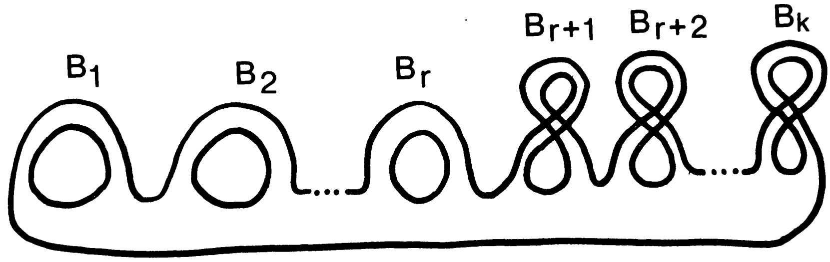


FIG. 11

Proof. We may assume (by using a regular homotopy) that f embeds the disk part of M in standard fashion and that the bands are immersed disjointly, with either one twist or no twist. If M is represented as in Figure 11 with bands $B_1, B_2, \ldots, B_k$, then we may assume that $B_1, \ldots, B_r$ have no twist, while $B_{r+1}, \ldots, B_k$ each have a single twist. This is accomplished by using a sequence of permutations as in Figure 12A. There are $(k+1)$ boundary curves $C_0, C_1, \ldots, C_k$. The curve C_0 is the outer boundary. Hence $B(f) = (k-r) + N(C_0)$. Since $N(C_0) = 0$ or 1, this presentation is not quite canonical. However, if $k - r \equiv 0 \pmod 2$ then $N(C_0) = 0$, and if $k - r \equiv 1 \pmod 2$ then $N(C_0) = 1$ (since $B(f)$ is even). If k is odd and $B(f) = k+1$, there is no ambiguity. Otherwise, there are two immersions of this form corresponding to each value of $B(f)$, one with $N(C_0) = 0$ and one with $N(C_0) = 1$. They are image homotopic via the "handle-sliding" operation illustrated in Figure 12 for $k = 2$ and $k = 4$.

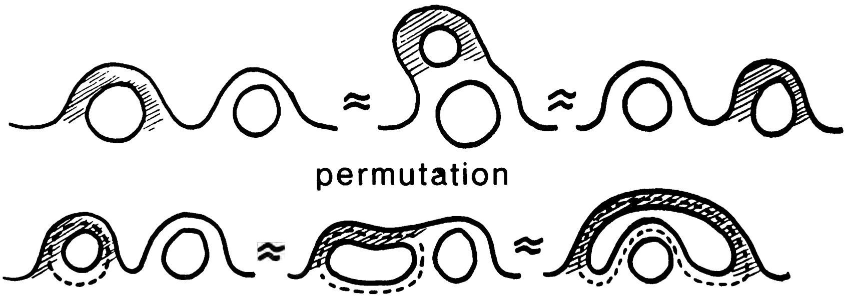


FIG. 12A

In the general case we slide the right hand end of B_1 across all other bands, cancel some twists using the mod-2 character of immersions in S^2, and proceed as in Figure 12. This completes the proof.

6. Curves, homology, and quadratic forms. If M is an arbitrary surface and $f: M \to S^2$ an immersion, then the boundary invariant is insufficient to determine the image homotopy class of f. We

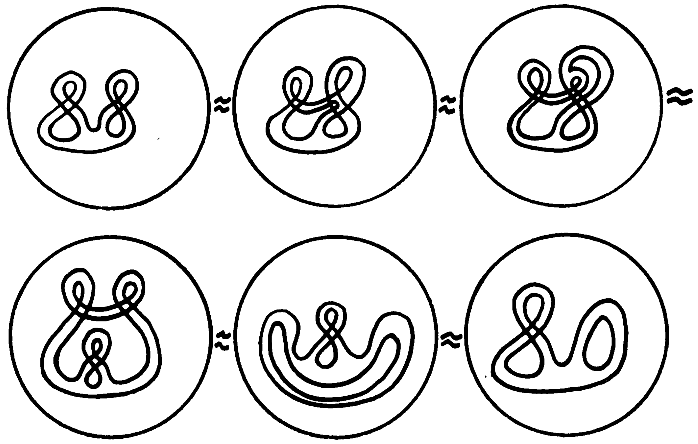


FIG. 12B

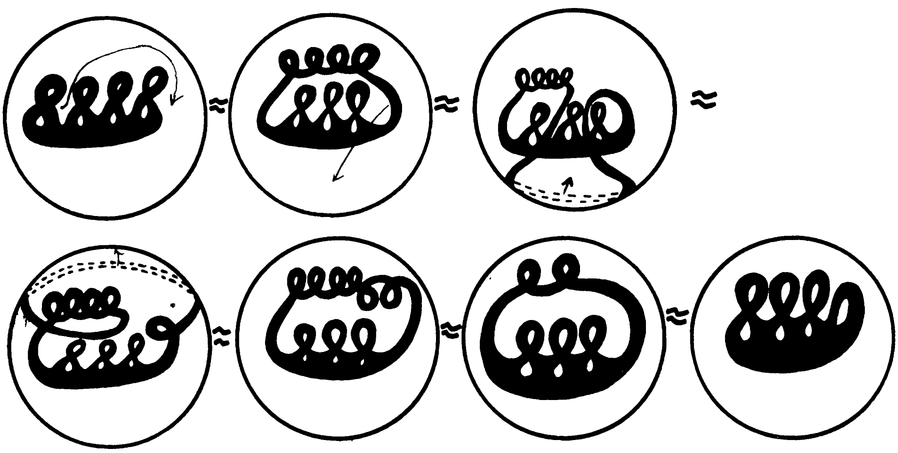


FIG. 12C

shall see that $N(f): \mathscr{C}(M) \to Z_2$ contains all of the extra information that is needed. However, $\mathscr{C}(M)$ is a very large collection of curves; some simplification is called for.

Suppose $\alpha, \beta \in \mathscr{C}(M)$. It may happen that α is regularly homotopic to β. That is, there may be a map $F: S^1 \times [0,1] \to M$ such that each F_t, $0 \leqq t \leqq 1$, is an immersion and $F_0 = \alpha$, $F_1 = \beta$. Under these circumstances $f \circ \alpha$ and $f \circ \beta$ are regularly homotopic immersions of S^1 to S^2. (The regular homotopy is $f \circ F: S^1 \times [0,1] \to S^2$.) Thus $N(f)(\alpha) = N(f)(\beta)$.

DEFINITION 6.1. Let $\bar{\mathscr{C}}(M)$ denote the collection of regular homotopy classes of elements of $\mathscr{C}(M)$. Given $\alpha \in \mathscr{C}(M)$, let $\bar{\alpha}$ denote its regular homotopy class in $\bar{\mathscr{C}}(M)$. Then we may define $\bar{N}(f): \bar{\mathscr{C}}(M) \to Z_2$ by $\bar{N}(f)(\bar{\alpha}) = N(f)(\alpha)$. The remarks above assure us that this definition makes sense.

Now $\bar{\mathscr{C}}(M)$ has extra structure which can be exploited. Given α and $\beta \in \mathscr{C}(M)$, we can deform each of them by a regular homotopy so that $\alpha(S^1)$ and $\beta(S^1)$ intersect normally. Assuming that α and β intersect normally, let $\alpha \cdot \beta \in Z_2$ denote the mod-2 residue class of their number of mutual intersections. This intersection number in Z_2 depends only on the regular homotopy classes of the curves. We obtain a pairing $\cdot: \bar{\mathscr{C}}(M) \times \bar{\mathscr{C}}(M) \to Z_2$ by setting $\bar{\alpha} \cdot \bar{\beta} = \alpha \cdot \beta$ where α and β are chosen to intersect normally.

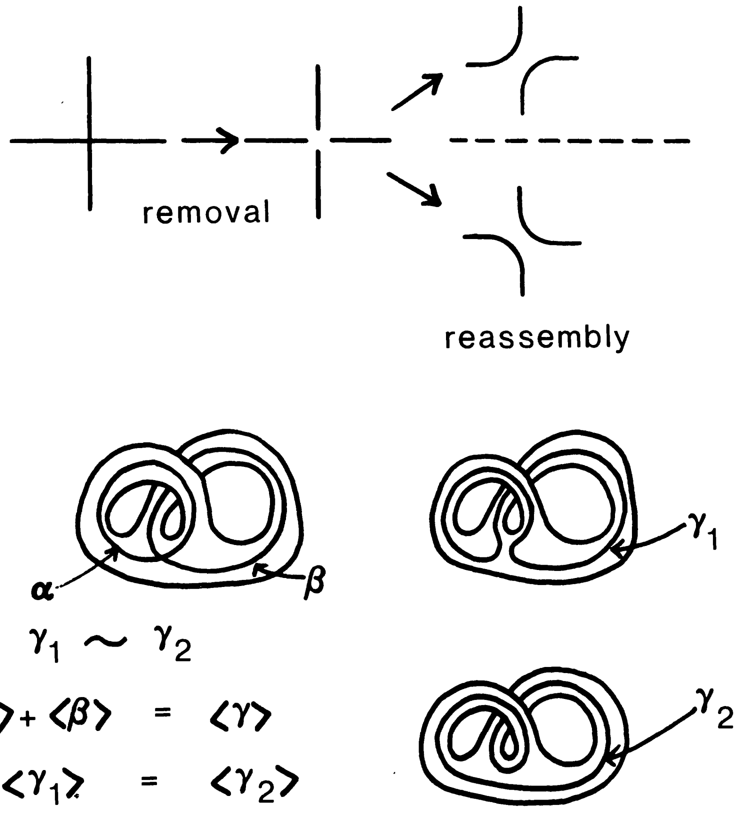


FIG. 13

It is also possible to add elements of $\mathscr{C}(M)$. This involves removing intersection points. Suppose two arcs intersect normally as in Figure 13. There are two ways to remove the intersection point and reassemble the remainder into two non-intersecting arcs (see Figure 13). Given curves $\alpha, \beta \in \mathscr{C}(M)$ intersecting normally, we may systematically remove all the intersection points and then reassemble to obtain a connected closed curve with no self-intersections. There are many ways to do this. Call the set of curves obtained in this way $\alpha \oplus \beta$.

Thus we have a procedure for addition:

(1) Choose $\alpha, \beta \in \mathscr{C}(M)$.

(2) Find $\alpha', \beta' \in \mathscr{C}(M)$ so that α' and β' intersect normally, and $\alpha' \approx \alpha$, $\beta' \approx \beta$.

(3) Remove all intersection points of α' and β'.

(4) Reassemble to form an embedded curve γ.

(5) Then $\gamma \in \alpha \oplus \beta$.

There is an important relationship between this addition process and $N(f)$.

LEMMA 6.2. *Let $f: M \to S^2$ be an immersion. Let $\alpha, \beta \in \mathscr{C}(M)$ be two normally intersecting curves. Then, for any choice of γ in $\alpha \oplus \beta$,*

$$N(f)(\gamma) = N(f)(\alpha) + N(f)(\beta) + \alpha \cdot \beta.$$

Proof. We may assume that $\alpha(S^1) \cap \beta(S^1) \neq \varnothing$ since this can be arranged by a regular homotopy. By the same reasoning, we may assume that $f \circ \alpha(S^1)$ and $f \circ \beta(S^1)$ have normal self and mutual intersections. Now $N(f)(\gamma)$ equals the total mod-2 number of intersections of $f \circ \alpha(S^1)$ and $f \circ \beta(S^1)$ (self and mutual) minus those mutual intersections which already occur on M. Since the total mutual

intersection of two curves on S^2 is even, we conclude that

$$N(f)(\gamma) = N(f)(\alpha) + N(f)(\beta) + \alpha \cdot \beta.$$

This lemma suggests that it would be advantageous to place an equivalence relation on $\mathscr{C}(M)$ so that both sums and the mod-2 intersection number are well-defined on the equivalence classes. Such an equivalence relation must include regular homotopy; it should also make all elements of the set $\gamma_1 \oplus \gamma_2$ equivalent when γ_1 and γ_2 are normally intersecting curves on the surface. Thus we make the following definition:

DEFINITION 6.3. We say that two curves $\alpha, \beta \in \mathscr{C}(M)$ are *homologous* ($\alpha \sim \beta$) if one may be obtained from the other by a finite sequence of elementary homologies. An elementary homology may be of two types:

(1) $\alpha \sim \beta$ if α and β are regularly homotopic on M ($\bar{\alpha} = \bar{\beta} \in \bar{\mathscr{C}}(M)$).

(2) $\alpha \sim \beta$ if there are curves $\gamma_1, \gamma_2 \in \mathscr{C}(M)$, intersecting normally, so that α and β are each members of $\gamma_1 \oplus \gamma_2$.

Let $\mathscr{H}(M)$ denote the set of homology classes of curves in $\mathscr{C}(M)$. If $\alpha \in \mathscr{C}(M)$, let $\langle\alpha\rangle \in \mathscr{H}(M)$ be its homology class.

Given $\langle\alpha\rangle, \langle\beta\rangle \in \mathscr{H}(M)$, we may assume that the representatives α and β have non-empty, normal intersection. (Since the surface is connected, it is always possible to change two curves by a regular homotopy so that they have non-empty intersection.) We therefore define the sum by the equation $\langle\alpha\rangle + \langle\beta\rangle = \langle\gamma\rangle$ for any γ in $\alpha \oplus \beta$. The intersection pairing is defined by the equation $\langle\alpha\rangle \cdot \langle\beta\rangle = \alpha \cdot \beta$. In fact $\mathscr{H}(M)$ becomes a group. The identity element, 0, is represented by any small curve about a point on M. Note that it follows from Lemma 6.2 that $N(f)$ is well-defined on $\mathscr{H}(M)$ and that

$$N(f)(\langle\alpha\rangle + \langle\beta\rangle) = N(f)(\langle\alpha\rangle) + N(f)(\langle\beta\rangle) + \langle\alpha\rangle \cdot \langle\beta\rangle.$$

(We set $N(f)(\langle\alpha\rangle) = N(f \circ \alpha)$.)

It is not hard to verify the next lemma:

LEMMA 6.4. *If M is any surface, then $\mathscr{H}(M)$ is an abelian group such that every element has order* 2. *Thus we may regard $\mathscr{H}(M)$ as a Z_2-vector space. If M has normal form as in Figure 6 then $\mathscr{H}(M)$ has Z_2-basis $\mathscr{B} = \{a_1, a'_1, \cdots, a_r, a'_r, b_1, \cdots, b_s\}$ where each of these classes represents a curve which traverses a single band on M.*

The intersection pairing $\cdot : \mathscr{H}(M) \times \mathscr{H}(M) \to Z_2$ is bilinear and symmetric, and $x \cdot x = 0$ for all $x \in \mathscr{H}(M)$. In the basis $\mathscr{B}$,

$$\begin{aligned} &a_i \cdot a'_i = 1 \\ &a_i \cdot a_j = a_i \cdot a'_j = 0, \qquad i \neq j \\ &a_i \cdot b_j = a'_i \cdot b_j = 0, \qquad \text{any } i, j. \end{aligned}$$

(Compare with Figure 6.)

In fact, $\mathscr{H}(M) \simeq H_1(M; Z_2)$, the usual first homology group of M with Z_2 coefficients. (See **[2]**, pp. 92–94.)

We are now in a position to reformulate Lemma 5.2.

DEFINITION 6.5. Let V be a finite dimensional Z_2 vector space and $\cdot : V \times V \to Z_2$ a symmetric, bilinear form such that $x \cdot x = 0$ for all $x \in V$. We say that a function $q : V \to Z_2$ is a *mod-2 quadratic form* associated with the pairing $\cdot$ if

$$q(x + y) = q(x) + q(y) + x \cdot y$$

for all $x, y \in V$.

COROLLARY 6.6. *Let $f: M \to S^2$ be an immersion. Define $q(f): \mathcal{H}(M) \to Z_2$ by $q(f)\langle\alpha\rangle = N(f)(\alpha)$. Then $q(f)$ is a mod-2 quadratic form associated with the intersection pairing on $\mathcal{H}(M)$.*

Note that, by 6.4, we can actually compute $q(f)$ by finding the mod-2 degrees of each band-curve on M. How does $q(f)$ behave under an image homotopy of f? Obviously it remains unchanged under regular homotopy of f. The next lemma shows what happens when the image homotopy involves a diffeomorphism:

LEMMA 6.7. *If $h: M \to M$ is a diffeomorphism, then h induces a vector space isomorphism $h_*: \mathcal{H}(M) \to \mathcal{H}(M)$ and $(h_*a)\cdot(h_*b) = a \cdot b$ for all $a, b \in \mathcal{H}(M)$.*

Proof. Define $h_*\langle\alpha\rangle = \langle h \circ \alpha\rangle$. The rest is easy to check.

Now suppose that $g = f \circ h$ where h is a diffeomorphism of M. Then $q(g)\langle\alpha\rangle = N(f \circ h \circ \alpha) = q(f)\langle h \circ \alpha\rangle = q(f) \circ h_*\langle\alpha\rangle$. Thus $q(g) = q(f) \circ h_*$.

DEFINITION 6.8. Let $q, q': V \to Z_2$ be two quadratic forms. One says that q is isomorphic to q' ($q \simeq q'$) if there is a vector space isomorphism $T: V \to V$ such that $q' = q \circ T$.

COROLLARY 6.9. *If f and $g: M \to S^2$ are image homotopic immersions, then the quadratic forms $q(f)$ and $q(g)$ are isomorphic.*

We can now prove that the punctured torus immersions T_0 and T_1 are not image homotopic. Let $\phi_0 = q(T_0)$ and $\phi_1 = q(T_1)$. Then $V = \mathcal{H}(M) \simeq Z_2 \oplus Z_2$ with basis $\{a, b\}$ so that $\phi_0(a) = \phi_0(b) = 0$, $\phi_1(a) = \phi_1(b) = 1$. Hence $\phi_0(a+b) = a \cdot b = 1$ and $\phi_1(a+b) = 1+1+1 = 1$. Therefore ϕ_1 takes a majority of elements of V to 1 while ϕ_0 takes a majority of elements of V to 0. If $\phi_1 = \phi_0 \circ f$ where $f: V \to V$ is an isomorphism, then ϕ_1 would also take a majority to 0. Therefore ϕ_0 and ϕ_1 are not isomorphic, and hence T_0 and T_1 are not image homotopic.

On the other hand, our pictures of homotopies give rise to specific isomorphisms. Consider the handle-sliding operation of Figure 12. If b denotes the curve (dotted) on the band being moved, while a is the curve on the stationary band, then our picture shows that $a \mapsto a$ while $b \mapsto a + b$ under the image homotopy.

For example, Figure 8 shows an image homotopy of T_0 to T_{01}. We know that $q(T_0) = \phi_0$; let $\phi = q(T_{01})$. Then the isomorphism $h_*: V \to V$ is given by the equations $h_*(a) = a$ and $h_*(b) = a + b$.

The association of immersions and quadratic forms is a double-edged tool. It will let us determine the structure of both immersions and quadratic forms.

7. Basic homotopies and isomorphisms. Immersions may be combined. The beginning of Figure 14 illustrates the result of joining two copies of T_0 together by an untwisted band. This is called the connected sum. Given any two immersions A and B, their connected sum, $A \# B$, is formed in the same way. If one of the surfaces has more than one boundary component, then it is possible to form different immersions by joining along different boundary components. For example, Figure 10 illustrates two ways of taking the connected sum of Λ_1, Λ_1, and Λ_0. Nevertheless, the symbol $A \# B$ will denote any one of the possible connected sums of A and B.

The algebraic analogue of connected sum is the direct sum of quadratic forms.

DEFINITION 7.1. Let $q: V \to Z_2$ and $q': V' \to Z_2$ be mod-2 quadratic forms. Define the direct sum, $q \oplus q': V \oplus V' \to Z_2$ by the formula $q \oplus q'(r, r') = q(r) + q'(r')$ for $r \in V$ and $r' \in V'$. It is easy to see that $q \oplus q'$ is a quadratic form.

If A and B are immersions of surfaces M and N, then $A \# B$ is an immersion of $M \# N$ and $\mathcal{H}(M \# N) \simeq \mathcal{H}(M) \oplus \mathcal{H}(N)$. It is clear that $q(A \# B) \simeq q(A) \oplus q(B)$.

For example, let Λ_0 and Λ_1 be the untwisted and single-twist annulus immersions. Let η_0 and η_1 be the mod-2 quadratic forms on $V = Z_2$ defined by setting $\eta_0(1) = 0$ and $\eta_1(1) = 1$. Then $q(\Lambda_0) = \eta_0$ and

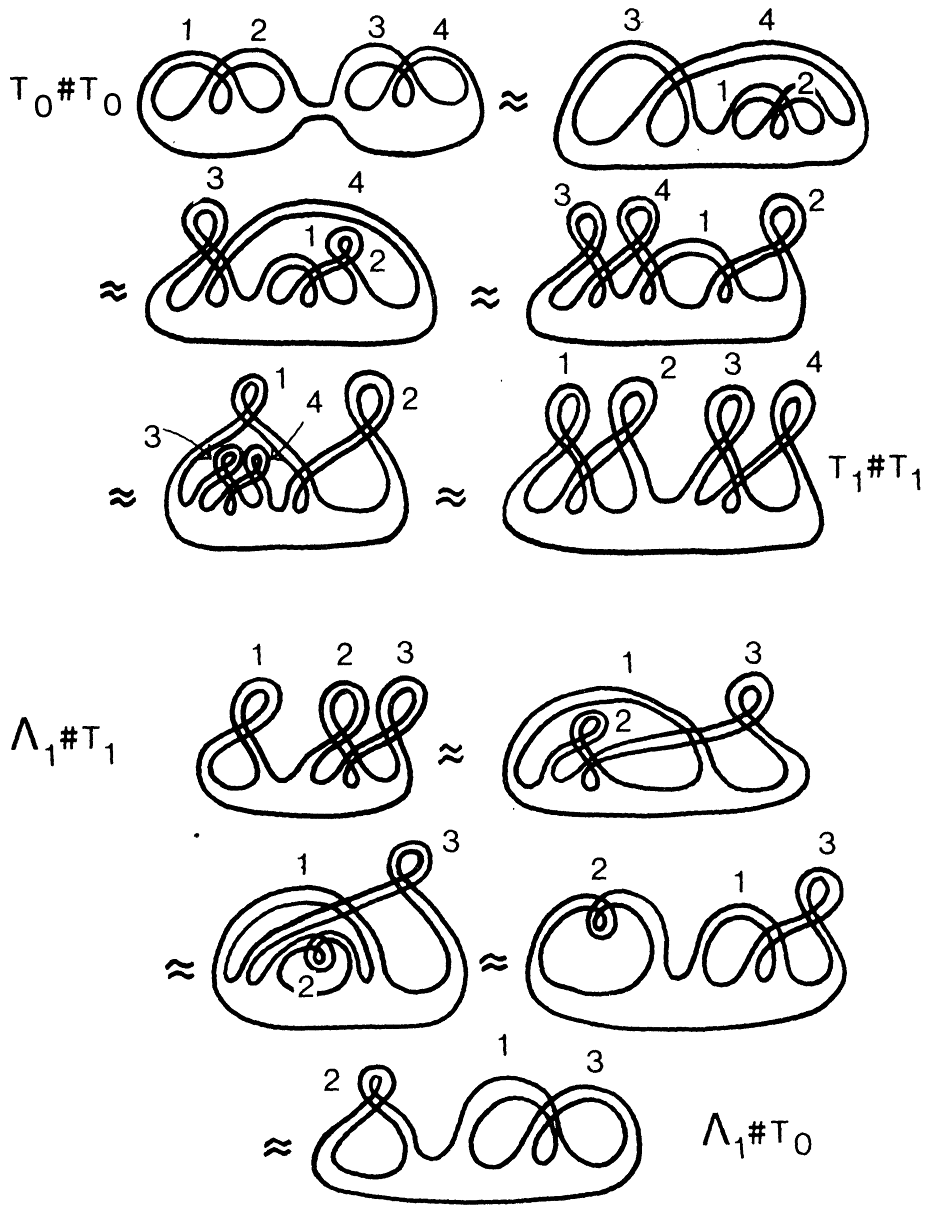


FIG. 14

$q(\Lambda_1) = \eta_1$. Figure 12B may be interpreted as $\Lambda_0 \# \Lambda_1 \approx \Lambda_1 \# \Lambda_1$. Hence $\eta_0 \oplus \eta_1 \simeq \eta_1 \oplus \eta_1$. Since direct sum of forms is well-defined, commutative, and associative, this implies that $\eta_1 \oplus \eta_1 \oplus \eta_1 \simeq \eta_0 \oplus \eta_0 \oplus \eta_1$. This type of reduction occurs for forms but not for immersions (the boundary invariant gets in the way).

Certain basic homotopies lead to fundamental isomorphisms of quadratic forms:

THEOREM 7.2. (a) $T_0 \# T_0 \approx T_1 \# T_1$; (b) $\Lambda_1 \# T_1 \approx \Lambda_1 \# T_0$; (c) $\Lambda_1 \# \Lambda_1 \approx \Lambda_0 \# \Lambda_1$.

Hence, (a′) $\phi_0 \oplus \phi_0 \simeq \phi_1 \oplus \phi_1$, (*recall that* $\phi_0 = q(T_0)$, $\phi_1 = q(T_1)$); (b′) $\eta_1 \oplus \phi_1 \simeq \eta_1 \oplus \phi_0$; (c′) $\eta_1 \oplus \eta_1 \simeq \eta_0 \oplus \eta_1$.

Proof. The proof involves handle sliding as illustrated in Figure 14. For (a) the bands are labelled 1, 2, 3, 4. The first step involves sliding the $1-2$ group over 3. Then 3 is slid around 1, 2 and 4; and 2 is slid around 1. Note that in sliding 3 past the $1-2$ group one slides 3 along 1, then around 2, then around 1, and finally across 2. This adds two twists to band 3. Sliding around 4 adds another twist. Thus in the second step, band 3 acquires 3 twists. Since twists cancel in pairs on S^2 (by the global swing-around of Figure 3) we have illustrated band 3 with the single resulting twist. This mod-2 twist arithmetic goes on throughout most of the rest of the deformation.

In the third step, 4 slides over 2, and acquires a twist. Then 1 slides over 3 acquiring a twist. Finally the $3-4$ group slides out across 2 and we have $T_1 \# T_1$.

For (b) the procedure is similar. Slide 1 across 2. Slide the left end of 2 around 1. Slide 2 out. Slide 3 around 1.

We have discussed (c) in the remarks prior to the theorem. This completes the proof.

The algebra isomorphisms (a′), (b′), (c′) deserve algebraic proofs. It is not hard to ferret out proofs from our homotopies by following a homology basis throughout the deformations. For example, let $\psi = q(T_0 \# T_0)$ and $\psi' = q(T_1 \# T_1)$. Let V have basis a_1, a_2, a_3, a_4 where a_i is the curve corresponding to the band labelled i. Then $\psi(a_1) = \psi(a_2) = \psi(a_3) = \psi(a_4) = 0$ while $\psi'(a_1) = \cdots = \psi'(a_4) = 1$. Here are the transformations corresponding to each step in the deformation for part (a) (refer to Figure 14).

(1) identity

(2) $a_1 \mapsto a_1$
$a_2 \mapsto a_1 + a_2$
$a_3 \mapsto a_3 + a_4$
$a_4 \mapsto a_4$

(3) $a_i \mapsto a_i \quad i = 1, 2, 3$
$a_4 \mapsto a_2 + a_4$

(4) $a_1 \mapsto a_1 + a_3$
$a_i \mapsto a_i, \quad i = 2, 3, 4$

(5) identity.

These are obtained by repeatedly using the handle sliding basis-change discussed at the end of the last section. Note that sliding across a torus-group (like 1–2) induces the identity transformation since one must slide past each band in the group twice.

The mapping $h_*: V \to V$ is the composite of these five maps. We find

$$h_*(a_1) = a_1 + a_3$$
$$h_*(a_2) = a_1 + a_2 + a_3$$
$$h_*(a_3) = a_2 + a_3 + a_4$$
$$h_*(a_4) = a_2 + a_4.$$

It is easy to check that $\psi' \circ h_* = \psi$. This gives an algebraic proof that $\phi_0 \oplus \phi_0 \simeq \phi_1 \oplus \phi_1$.

The isomorphisms of Theorem 7.2 are the key to the classification of mod-2 quadratic forms! Here is the algebra in its own right:

Let $q: V \to Z_2$ be an arbitrary mod-2 quadratic form with associated pairing $\cdot : V \times V \to Z_2$. When $q = q(f)$ for an immersion $f: M \to S^2$ we know that V has a basis $\mathcal{B}$ as in 5.4. In fact one can always find a basis with these intersection properties for any symmetric bilinear pairing on V such that $x \cdot x = 0$ for all $x \in V$. The proof (algebraic proof) is a standard exercise in linear algebra. It then follows, just as in the geometric case, that any form $q: V \to Z_2$ is a direct sum involving the forms ϕ_0, ϕ_1, η_0 and η_1. Such a direct sum reduces to one of the following types by using (a′), (b′), and (c′).

(i) $\phi_1 \oplus l\phi_0 \oplus (s+1)\eta_0$

(ii) $(l+1)\phi_0 \oplus (s+1)\eta_0$

(iii) $(l+1)\phi_0 \oplus s\eta_0 \oplus \eta_1$.

Here $l\phi = \phi \oplus \phi \oplus \cdots \oplus \phi$ (l-factors).

These types are distinct. To see this, first characterize the subspace of V which supports the η_0 and η_1 factors. This is the *radical* of V, denoted Rad V. It is the subspace

$$\text{Rad } V = \{x \in V \mid x \cdot y = 0, \quad \text{for all } y \text{ in } V\}.$$

Thus, forms of type (i) or (ii) have $q|\mathrm{Rad}\,V \equiv 0$. If this is the case, then we obtain $\bar{q}: V/\mathrm{Rad}\,V \to Z_2$ and $\bar{q}$ is *non-degenerate.* That is, the matrix of the bilinear form for $\bar{q}$ is non-singular. The form $\bar{q}$ is isomorphic to $\phi_1 \oplus l\phi_0$ or to $(l+1)\phi_0$. In the first case, a majority of the elements of $V/\mathrm{Rad}\,V$ go to 1, while in the second case a majority go to 0. We may therefore define the *Arf invariant* $c(q) = 1$ or 0 according to this majority vote by $\bar{q}$.

In case (iii) $q|\mathrm{Rad}\,V \not\equiv 0$. But forms of this type are characterized by dim(V) and dim (Rad V).

This completes the classification of mod-2 quadratic forms. In the next section we apply our results and classify immersions.

REMARK. It is useful at this point to survey the correspondence that we have obtained. If V is a vector space over a field F and $\langle\ ,\ \rangle: V \times V \to F$ is a symmetric bilinear form then, if $1+1 \neq 0$ in F, we may define a quadratic form $q: V \to F$ by $q(x) = \frac{1}{2}\langle x, x\rangle$. Note that we now have

$$q(x+y) = \tfrac{1}{2}\langle x+y, x+y\rangle = \tfrac{1}{2}(\langle x,x\rangle + \langle y,y\rangle + 2\langle x,y\rangle) = q(x) + q(y) + \langle x, y\rangle.$$

Thus q is a quadratic form associated with $\langle\ ,\ \rangle$. Conversely, if q is quadratic form then we obtain a corresponding bilinear form. This correspondence breaks down when $1+1=0$ and there may be many quadratic forms associated with a given bilinear form. Just so in our geometry we take one surface (and its homology intersection form) and we find many immersions of this surface into S^2 (and their quadratic forms). For a given surface, each quadratic form is associated with one given bilinear form (the intersection form). Since image homotopy of immersions implies isomorphism of the corresponding mod-2 quadratic forms, it has been possible to give a geometric version of the theory of mod-2 forms.

8. Classification of immersions. We are now prepared to complete the classification of immersed surfaces.

THEOREM 8.1. *Let f and g be orientation preserving immersions of a surface with boundary M into the two-sphere, S^2. Then f is image homotopic to g if and only if the boundary invariants of f and g agree and the quadratic forms are isomorphic. That is,*

$$\text{(a)}\quad B(f) = B(g) \quad \textit{and} \quad \text{(b)}\quad q(f) \simeq q(g).$$

REMARK. If $B(f) = B(g) = 0$, then $q(f)$ and $q(g)$ have Arf invariants (see section 7) and we may replace (b) by

$$\text{(b}'\text{)}\quad c(q(f)) = c(q(g)).$$

If $B(f) = B(g) \neq 0$, then the quadratic forms are of type (iii) and hence are classified by the dimensions of their radicals.

Proof. As in the proof of 5.3, we may assume that M is in normal form as in Figure 6 (a disk with attached bands), that f embeds the disk, and that the bands corresponding to different pairs $\{a_i, a'_i\}$ or singlets $\{b_j\}$ do not intersect in the image. In other words, f is a connected sum involving T_0, T_1, Λ_0 and Λ_1 with all the copies of T_0 and T_1 appearing on a single boundary component. Hence, by using 7.2 we may write, f is image homotopic to one of the following forms:

(i) $f \approx T_1 \# kT_0 \# l\Lambda_0$

(ii) $f \approx kT_0 \# l\Lambda_0$

(iii) $f \approx kT_0 \# l\Lambda_0 \# s\Lambda_1,\ s \neq 0.$

Here the connected sum is the specific one arising from the normal form for M. Immersions of type (i) and (ii) have $B(f) = 0$ and are distinguished by the quadratic form, as we have seen. An immersion of type (iii) is clearly distinguished by $B(f)$ (which tells how many Λ_1's appear) and $q(f)$ (which tells how many Λ_0's and Λ_1's appear). This completes the proof.

Note the close parallel with the classification of quadratic forms. Boundary difficulties, as in section 5, prevent the correspondence from being perfect. This was to be expected since there are no

distinguished elements in V for an arbitrary form $q: V \to Z_2$, while $\mathcal{H}(M)$ contains the homology classes of the boundary curves.

If M has a single boundary component then the quadratic form has no radical. Such forms are sums of ϕ_0 and ϕ_1. Theorem 8.1 implies that isomorphism classes of non-degenerate (Rad = 0) forms are in 1–1 correspondence with image homotopy classes of immersions with a single boundary component.

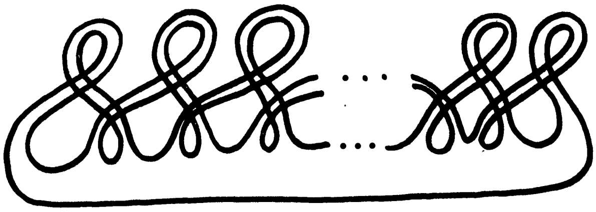


FIG. 15

Here is an exercise. Let N_k be the immersion pictured in Figure 15. It has k-bands. Reduce it to normal form by handle-sliding. What does this say about the corresponding quadratic forms as a function of $k = 1, 2, 3, 4, \ldots$?

Partially supported by NSF grants: MPS73–08855–A02 and GP–38582(X1), MPS73–08732–A02.

References

1. W. Browder, Surgery on Simply Connected Manifolds, Springer-Verlag, New York, 1972.
2. M. Greenberg, Lectures on Algebraic Topology, Benjamin, New York, 1967.
3. F. Hirzebruch, W. D. Neumann, and S. S. Koh, Differentiable Manifolds and Quadratic Forms, Marcel Dekker, New York, 1971.
4. W. Massey, Algebraic Topology: An Introduction, Harcourt Brace and World, New York, 1967.
5. J. Milnor, D. Husemoller, Symmetric Bilinear Forms, Springer-Verlag, New York, 1973.
6. C. P. Rourke and D. P. Sullivan, On the Kervaire obstruction, Ann. of Math., 94 (1971) 397–413.
7. H. Whitney, On regular closed curves in the plane, Comp. Math., 4 (1937) 276–284.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MI 48109.
DEPARTMENT OF MATHEMATICS, BROWN UNIVERSITY, PROVIDENCE, RI 02912.

THE FIFTH U.S.A. MATHEMATICAL OLYMPIAD

S. L. GREITZER

The Fifth U.S.A. Mathematical Olympiad was held on May 4, 1976. As before, it consisted of five power questions requiring mathematical ingenuity as well as competence and knowledge of subject matter. The problems will be found at the end of this article.

As in previous years, most students were selected on the basis of their performance on the Annual High School Mathematics Examination. Several students from Michigan and Wisconsin were invited to participate. These states do not, as a rule, participate in the Annual High School Mathematics Examination, but have their own contests. In all, 96 students were asked to participate, and 94 finally took part. One student did not reply to the invitation, and one student withdrew suddenly.

For reasons connected with the International Mathematical Olympiad, it was imperative that all papers be graded by May 14. This was done — in fact, the last paper arrived while the papers were being graded. The grading committee consisted of Professors Michael Aissen, John Bender, Richard Bumby, Harry Gonshor, Phil Guza, Ben Muckenhoupt, Barbara Osofsky, Sol Leader and Hyman Zimmerberg, all of Rutgers University Department of Mathematics. The top thirty papers were then regraded by Professors Aissen, Bender and Greitzer. The Olympiad Committee acknowledges with thanks the help provided by this grading committee.

Many individuals and organizations contributed help and support that helped the whole process run smoothly. The contest problems were prepared by a committee consisting of Murray Klamkin, C. C. Rousseau and Tom Griffiths. The Annual High School Mathematics Examination Committee, consisting of R. Artino, A. Gaglione and N. Shell, provided the data needed for selection of participants. We wish to express our gratitude for their help and support.

The Sixth U.S.A. Mathematical Olympiad is scheduled for Tuesday, May 3, 1977. As before, participation will be by invitation only, mainly on the basis of achievement on the Annual High School Mathematics Examination.

The results of the Olympiad, in order of rank, were as follows:

*Kleiman, Mark	Stuyvesant H. S.	New York, N.Y.
*Stephanides, Adam	U. of Chicago Lab. School	Chicago, Ill.
Herdeg, Paul	Hamilton-Wenham Reg. H. S.	So. Hamilton, Mass.
# Dougherty, Randall	W. T. Woodson H. S.	Fairfax, Va.
# Knierim, Daniel	Hiram Johnson Sr. H. S.	Sacramento, Cal.
Mifflin, Richard	Stratford Sr. H. S.	Houston, Texas
Puckette, Miller	Sewanee Academy	Sewanee, Tenn.
Kelly, Reed	Stuyvesant H. S.	New York, N.Y.

(* and # indicate equal scores.)

Messrs. Kleiman, Stephanides and Herdeg submitted perfect papers. Either the Olympaid was somewhat easier or, more probably, the better results were due to the fact that both Kleiman and Herdeg had attended the Training Session in 1975, at which the team we sent to the International Mathematical Olympiad received special instruction.

For purposes of analysis of results, we present a table of scores on the High School Mathematics Examination and the related scores on the Olympiad:

Olympiad / H. S. Exam	0	1–10	11–20	21–30	31–40	41–50	51–60	61–70	71–80	81–90	91–100	101–110
135–144												1
125–134	1											2
115–124	1	1			2	1						
105–114		2	3	4	3	1	2	1				
95–104		17	7	8	3	4	1	1				
85–94	1	5	3	1	2	3	4	1	1			

This table does not include three students from Wisconsin and two from Michigan. For these, the highest score was 58, the lowest score 10.

First, there is almost no correlation between the scores on the two tests. Second, there were only seven students who did not receive a flat zero on some problem. Six of these students are among those listed above. Third, 71 students were able to do something with the first problem, which required ingenuity only, whereas many more fell down on the other problems which required some knowledge of subject matter.

The table below gives the number of students attaining various grades on each of the problems on the Olympiad:

Score \ No.	1	2	3	4	5
21–25	1	4	0	3	4
16–20	34	16	14	5	18
11–15	8	5	3	0	0
6–10	6	2	3	6	1
1–5	21	18	24	52	4
0	24	49	50	28	67

Now it is quite possible that the very large number of low grades for Problem 5 is due to the possibility that students never managed to reach this last problem. However, even those who did reach it did not do too well. Only 22 students did fairly well or got a full score.

The five problems can be classified as follows: No. 1 required merely mathematical ingenuity. A large number of students did well on it. No. 2 was a problem in plane geometry. Few students got this one. No. 4 was a problem in solid geometry and inequalities. This was apparently the hardest. No. 3 was a problem in number theory, requiring only an understanding of parity relations. Students found this difficult. Finally, No. 5 was a problem in algebra, and the students found this difficult.

To this writer, it appears that we have students with mathematical ingenuity — witness the results on Problem 1 — but that they have not yet learned such basics as geometry and algebra — witness Problems 2, 4, 5. We call these subjects "basic" because students cannot learn further mathematics, not even calculus, unless they know these basics. We suggest that these subjects be made part of the curriculum in the senior year of the secondary school.

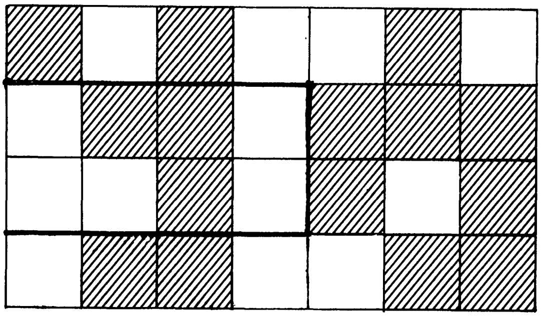
The finalists were honored at ceremonial excercises in Washington, D.C. on June 14 and 15. There was an awards ceremony at the National Academy of Sciences, a reception in the John Quincy Adams Room at the Department of State, and dinner in the Thomas Jefferson Room. As in preceding years, these ceremonies were made possible through the generosity of International Business Machines. Among the awards to the students were engraved silver trays from IBM, an HP–45 Calculator for each student from the Hewlett-Packard Company, and books from the Mathematical Association of America and the National Council of Teachers of Mathematics. (Because of health reasons, Adam Stephanides was unable to attend.)

The team representing the U.S.A. in Austria consisted of seven of the eight top scorers (Adam Stephanides could not come) plus an eighth student who had done well and had attended the training session the previous year. The International Mathematical Olympiad was held in Lienz, in Austria.

The U.S.S.R. scored first with 250 points, Great Britain was second with 214 points, and the U.S.A. third with 188 points. Hungary and the German Democratic Republic scored seventh and eighth respectively. The training of the team took place at Annapolis, in the U.S. Naval Academy, and travel expenses to Austria were supplied by the Army Research Office. The Olympiad Committee wishes to express its deep gratitude for the support of the training and travel.

FIFTH U.S.A. MATHEMATICAL OLYMPIAD — MAY 4, 1976

1.



(a) Suppose that each square of a 4×7 chessboard, as shown above, is colored either black or white. Prove that with *any* such coloring, the board must contain a rectangle (formed by the horizontal and vertical lines of the board), such as the one outlined in the figure, whose four distinct corner squares are all of the same color.

(b) Exhibit a black-white coloring of a 4×6 board in which the four corner squares of every such rectangle as described above, are not the same color.

2. If A and B are fixed points on a given circle and XY is a variable diameter of the same circle, determine (with proof) the locus of the point of intersection of lines AX and BY. You may assume that AB is not a diameter.

3. Determine (with proof) all integral solutions of $a^2 + b^2 + c^2 = a^2b^2$.

4. If the sum of the lengths of the six edges of a trirectangular tetrahedron $PABC$ (i.e., $\measuredangle APB = \measuredangle BPC = \measuredangle CPA = 90°$) is S, determine (with proof) its maximum volume.

5. If $P(x)$, $Q(x)$, $R(x)$, and $S(x)$ are all polynomials such that

$$P(x^5) + xQ(x^5) + x^2R(x^5) = (x^4 + x^3 + x^2 + x + 1)S(x),$$

prove that $x - 1$ is a factor of $P(x)$.

DEPARTMENT OF MATHEMATICS, RUTGERS UNIVERSITY, NEW BRUNSWICK, NJ 08903.

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Laboratoire Calcul des Probabilités, Université de Paris T.56, 4 Place Jussieu, 75–230 Paris, France.

A CONSTRUCTIVE LOOK AT ORTHONORMAL BASES IN HILBERT SPACE

DOUGLAS S. BRIDGES

As has been pointed out by Bishop [1] and others, a great deal of classical and modern mathematics has little or no computational significance. In order to recover those portions of the subject which do have this significance, the constructive mathematician is obliged to re-examine existent mathematics in the light of his primary contention that the phrase 'there exists' should be interpreted strictly as 'there can be computed', rather than in the idealistic sense which pervades so much of current mathematical practice. In other words, proofs of 'existence' which proceed by deriving a contradiction from the assumption of non-existence are unacceptable within constructive mathematics, and must, where possible, be replaced by proofs of computability of the objects under discussion (or, commonly, of suitable approximations to these objects).

The purpose of this note is to make some constructive remarks on orthonormal bases and dimensionality in Hilbert space. We shall say that a normed linear space E is *finite dimensional* if there exist finitely many elements $e_1, \ldots, e_n$ of E, and linear functionals $\phi_1, \ldots, \phi_n$ on E such that

$$x = \sum_{k=1}^{n} \phi_k(x) e_k \qquad (x \in E),$$

$$\phi_j(e_k) = 0 \qquad (1 \leqq j, k \leqq n, j \neq k),$$

and each ϕ_k is bounded — that is, there exists (we can compute!) $c > 0$ such that $|\phi_k(x)| \leqq c\|x\|$ for each x in E. The number n of elements of the *basis* $\{e_1, \ldots, e_n\}$ of E is then called the *dimension* of E, and is independent of the basis in question. We also consider the trivial space $\{0\}$ to be of finite dimension 0.

Of particular importance to us is the fact that, when E is a finite dimensional subspace of a normed linear space F, then E is *located*, in the sense that

$$\operatorname{dist}(x, E) \equiv \inf\{\|x - y\| : y \in E\}$$

is computable for each x in F. (This is not always the case for more general subspaces E.) Moreover, if F is a Hilbert space, then $\operatorname{dist}(x, E) = \|x - Px\|$, where P is the projection of F on E (cf. [1], Ch. 9, Sections 1,4). Bearing this in mind, we say that the normed linear space E is *infinite dimensional* if, whenever V is a finite dimensional subspace of E, there exists x in E with $\operatorname{dist}(x, V) > 0$.

In the special case of a Hilbert space H, dimensionality can be approached through the alternative medium of orthonormal bases: a sequence $(a_n)_{n \geqq 1}$ in a (necessarily separable) Hilbert space H is an *orthonormal basis* if

(i) $(a_m, a_n) = 0$ for $m \neq n$ (where $(,)$ is the scalar product on H),
(ii) for each n, either $a_n = 0$ or $\|a_n\| = 1$,
(iii) each x in H has a unique representation of the form $\Sigma_{n=1}^{\infty} x_n a_n$, with x_n in C and $x_n = 0$ whenever $a_n = 0$.

Classically, an orthonormal basis will be infinite if and only if H is infinite dimensional; moreover, all terms of an orthonormal basis will have norm 1. However, in order to obtain an analogue of the

Gram-Schmidt process for constructing orthonormal bases for arbitrary separable Hilbert spaces, the constructive mathematician is forced to define all orthonormal bases to be infinite, and to allow some (possibly none) of the terms of such a basis to be 0. Note that if (a_n) is an orthonormal basis of H, and x, y elements of H, then

$$x = \sum_{n=1}^{\infty} (x, a_n)a_n$$

and

$$(x, y) = \sum_{n=1}^{\infty} (x, a_n)(a_n, y).$$

For the remainder of this paper, we shall take H as a separable complex Hilbert space, and (a_n) an orthonormal basis of H. Now, if there exists ν such that $a_n = 0$ for $n \geqq \nu$, then it is clear that H is finite dimensional. On the other hand, if H is infinite dimensional, then we can construct a strictly increasing sequence $(n_k)_{k\geqq 1}$ such that $\| a_{n_k} \| = 1$ for each k. To see this, we set $a_0 = 0$, $n_0 = 0$, suppose $n_0, \ldots, n_k$ to have been found so that $\| a_{n_j} \| = 1$ for $1 \leqq j \leqq k$, and let V be the finite dimensional subspace of H spanned by $\{a_0, \cdots, a_{n_k}\}$. With P the projection of H on V, and choosing in turn x in H with $0 < \delta \equiv \operatorname{dist}(x, V)$, and $N > n_k$ so that

$$\sum_{n=N+1}^{\infty} |(x, a_n)|^2 < \delta^2/2,$$

we have

$$\begin{aligned} 0 < \delta^2 = \| x - Px \|^2 &= \sum_{n_k+1}^{\infty} |(x, a_n)|^2 \\ &\leqq \sum_{n_k+1}^{N} |(x, a_n)|^2 + \delta^2/2, \end{aligned}$$

so that $\Sigma_{n_k+1}^{N} |(x, a_n)|^2 > \delta^2/2$, and therefore we can find n_{k+1} with $n_k + 1 \leqq n_{k+1} \leqq N$ and $|(x, a_{n_{k+1}})| > 0$ ([1], Ch. 2, Propn. 7). It now follows from the Cauchy–Schwarz inequality that

$$\| a_{n_{k+1}} \| \geqq \| x \|^{-1} |(x, a_{n_{k+1}})| > 0,$$

so that $\| a_{n_{k+1}} \| = 1$.

These remarks naturally raise the questions:

* *If H is finite dimensional, does there exist ν such that $a_n = 0$ for $n \geqq \nu$?*

** *If there exists a strictly increasing sequence (n_k) such that $\| a_{n_k} \| = 1$ for each k, is H infinite dimensional?*

The affirmative answers to both these questions will appear as consequences of the following

LEMMA. *If V is a locally compact subspace of H, then there exists ν such that* $\operatorname{dist}(a_n, V) > 0$ *whenever $n \geqq \nu$ and $\| a_n \| = 1$.*

Proof. We first remark that, in constructive analysis, a metric space is defined to be *compact* if it is totally bounded and complete; that a *locally compact* space is a nonvoid metric space in which each bounded subset is contained in a compact subset; that a locally compact subspace of a metric space is located; and that the unit ball of a locally compact normed linear space is compact ([1], Ch. 4, Propn. 13, and Ch. 9, Propn. 5). Thus there exist points $x_1, x_2, \ldots, x_m$ of the unit ball V_1 of V such that the open balls with centres x_k and radius $\frac{1}{2}$ cover V_1. For each k we have $\| x_k \|^2 = \Sigma_{n=1}^{\infty} |(x_k, a_n)|^2$; whence we may choose ν_k so that $|(x_k, a_n)| \leqq \frac{3}{8}$ for $n \geqq \nu_k$. With P the projection of H on V, $\nu \equiv \max(\nu_1, \ldots, \nu_m)$, and $n \geqq \nu$, we now suppose that $\| a_n \| = 1$. Then, choosing k so that $1 \leqq k \leqq m$ and $\| Pa_n - x_k \| < \frac{1}{2}$, we have

$$\|x_k - a_n\|^2 \geqq 1 - 2\mathrm{Re}(x_k, a_n) > \tfrac{1}{4},$$

whence

$$\operatorname{dist}(a_n, V) = \|a_n - Pa_n\| \geqq \|a_n - x_k\| - \|x_k - Pa_n\| > 0,$$

as we required.

The affirmative answer to question * is now seen as a trivial consequence of the case $V = H$ of this lemma and the fact that a finite dimensional normed linear space is locally compact. On the other hand, if there exists a strictly increasing sequence (n_k) such that $\|a_{n_k}\| = 1$ for each k, then for each given finite dimensional linear subspace V of H we can compute in turn ν as in the above lemma, and k with $n_k \geqq \nu$, to obtain $\operatorname{dist}(a_{n_k}, V) > 0$; whence H is infinite dimensional.

Concluding Remarks:

1. For each integer $n \geqq 1$, let a_n be the sequence with all terms other than the nth equal to 0, and the nth equal to 1 or 0 according as it is, or is not, true that $2k$ is a sum of two primes for $k = 1, 2, \ldots, n$. Let H be the set of all sequences $(x_k)_{k \geqq 1}$ in C such that $\Sigma_{k=1}^{\infty} |x_k|^2$ converges, and $x_k = 0$ when $a_k = (0)$. Then H, taken with the scalar product

$$((x_k), (y_k)) \equiv \sum_{k=1}^{\infty} x_k \bar{y}_k,$$

is a complex Hilbert space, and (a_n) is an orthonormal basis of H. For this particular space, it is easy to see that the following three propositions are equivalent:

a. Either H is finite dimensional or H is infinite dimensional.

b. Either we can compute ν such that $a_n = 0$ for $n \geqq \nu$, or we can construct a strictly increasing sequence (n_k) such that $\|a_{n_k}\| = 1$ for each k.

c. Either we can compute an integer ν such that 2ν is not a sum of two primes, or $2n$ is a sum of two primes for each integer $n \geqq 1$.

In other words, a *constructive* proof of either a or b would produce a finite, routine method for deciding the Goldbach Conjecture. As the latter could be replaced in the above by virtually any unsolved problem of number theory, it seems unlikely that a constructive proof of either a or b will ever materialise.

2. Another (immediate) consequence of the case $V = H$ of our lemma is the proposition

a locally compact Hilbert space is finite dimensional,

a result for which we known of no simpler constructive proof.

Reference

1. Errett Bishop, Foundations of Constructive Analysis, McGraw-Hill, New York, 1967.

WOLFSON COLLEGE, OXFORD, ENGLAND, AND
UNIVERSITY COLLEGE, BUCKINGHAM MK18 1EG, ENGLAND.

ON THE EXISTENCE OF LOCALLY RECURRENT FUNCTIONS

MICHAEL L. ENGQUIST

We follow Bush [1] in saying that a function f is locally recurrent at x_0 if every deleted neighborhood of x_0 contains an element x such that $f(x) = f(x_0)$. In this note we investigate how prevalent certain types of continuous locally recurrent functions are in the sense of category in $C[0, 1]$.

Bush [1], [2] has investigated the existence of various types of continuous locally recurrent functions. His methods, which are quite different from ours, are based on n-ary expansions of real numbers.

We let $C[0,1]$ denote the set of continuous real-valued functions on $[0,1]$ equipped with the topology of the uniform norm. The norm of $f \in C[0,1]$ will be denoted by $\|f\|$.

We recall that the Baire Category Theorem [3, p. 68] implies that for a set X of first category in $C[0,1]$, the complement X' is dense. Hence, to show a certain set of functions Y is nonempty by the Baire Category Theorem, one tries to show that $Y = X'$ where X is a set of first category. This is the strategy followed by Banach to show the existence of nowhere differentiable functions [3, p. 260]. Our Theorem 1 shows that this strategy will not work to show the existence of nonconstant functions in $C[0,1]$ which are everywhere locally recurrent; however, as Theorem 2 shows, it will work if we relax our requirements slightly.

THEOREM 1. *The set*

$$A = \{f: f(x) = \sup f \text{ holds for at least two values of } x\}$$

is a set of first category in $C[0,1]$.

Proof: We let

$$A_m = \{f: f(x) = \sup f \text{ at least at } x_1 \text{ and } x_2 \text{ satisfying } |x_1 - x_2| \geqq 1/m\}.$$

Suppose that $\{f_k\}_{k=1}^{\infty} \subseteq A_m$ and $f_k \to f$ as $k \to \infty$. To each f_k there correspond x_{1k} and x_{2k} satisfying the conditions which define A_m. By taking a subsequence of $\{f_k\}_{k=1}^{\infty}$ if necessary, we may assume that $x_{1k} \to \bar{x}_1$ and $x_{2k} \to \bar{x}_2$ as $k \to \infty$. Clearly, $|\bar{x}_1 - \bar{x}_2| \geqq 1/m$. Since $\{f_k\}_{k=1}^{\infty}$ is uniformly convergent, it is equicontinuous. It follows from equicontinuity and the convergence of $\{x_{1k}\}_{k=1}^{\infty}$ and $\{x_{2k}\}_{k=1}^{\infty}$ that for a given $\varepsilon > 0$ there exists K_1 such that $k \geqq K_1$ implies $|f_k(x_{ik}) - f_k(\bar{x}_i)| < \varepsilon/2$. Since $\{f_k\}_{k=1}^{\infty}$ converges, there exists K_2 such that $k \geqq K_2$ implies $|f_k(\bar{x}_i) - f(\bar{x}_i)| < \varepsilon/2$. An application of the triangle inequality yields that $k \geqq \max\{K_1, K_2\}$ implies $|f_k(x_{ik}) - f(\bar{x}_i)| < \varepsilon$. Thus it is possible to pass to the limit in the inequality $f_k(x_{ik}) \geqq f_k(x)$, $x \in [0,1]$, to obtain $f(\bar{x}_i) = \sup f$ for $i = 1, 2$. Hence A_m is closed.

To show A_m is nowhere dense, it suffices to note that for $f \in C[0,1]$ and $\varepsilon > 0$ the function g defined by $g(x) = f(x) + \max\{0, \varepsilon - |x - x_0|\}$ satisfies $\|f - g\| = \varepsilon$ and $g \in A'_m$. Since $A = \bigcup_{m=1}^{\infty} A_m$, the proof is finished.

If I is a closed subinterval of $[0,1]$, $f \in C[0,1]$ and $y \in f(I)$, then a **crossing interval** of f in I at y-level is defined to be an open interval in I containing points u_1 and u_2 such that $f(u_1) < y < f(u_2)$. We say that the **crossing index** of f on I at y-level (denoted by $\nu(f, I, y)$) is p in case f has p pairwise disjoint crossing intervals in I at y-level but f does not have $p+1$ pairwise disjoint crossing intervals in I at y-level.

We recall that the complement of a set of first category is called a residual set.

THEOREM 2. *The set B of functions in $C[0,1]$ which are not constant on any interval and which are locally recurrent everywhere except possibly at points where a local max or min occurs is a residual set.*

Proof: We let $\{I_m\}_{m=1}^{\infty}$ be the sequence of closed subintervals of $[0,1]$ having rational end points. For positive integers m and n we define

$$E_{mn} = \left\{f: \nu(f, I_m, y) \leqq 1 \text{ for some } y,\ \inf(f|I_m) + \frac{1}{n} \leqq y \leqq \sup(f|I_m) - \frac{1}{n}\right\}.$$

First, we verify that E_{mn} is closed. Suppose that $\{f_k\}_{k=1}^{\infty} \subseteq E_{mn}$ and $f_k \to f$ as $k \to \infty$. There exists $y_k \in f_k(I_m)$ such that

$$\inf(f_k|I_m) + \frac{1}{n} \leqq y_k \leqq \sup(f_k|I_m) - \frac{1}{n} \quad \text{and} \quad \nu(f_k, I_m, y_k) \leqq 1.$$

By taking a subsequence of $\{f_k\}_{k=1}^{\infty}$ if necessary, we may assume that $y_k \to y$ as $k \to \infty$. Clearly

$y \in f(I_m)$ and $\inf(f \mid I_m)+(1/n) \leqq y \leqq \sup(f \mid I_m)-(1/n)$. We claim that $\nu(f, I_m, y) \leqq 1$. Suppose not. Then there exist at least two disjoint crossing intervals of f in I_m at y-level. If J_i denotes the ith crossing interval containing u_{i1} and u_{i2} such that $f(u_{i1})<y<f(u_{i2})$ for $i=1,2$, then we set

$$\varepsilon=\frac{1}{2} \min \{|f(u_{ij})-y|: 1 \leqq i, j \leqq 2\}.$$

If K_1 is chosen so that $\|f_k-f\|<\varepsilon$ for $k \geqq K_1$ and K_2 is chosen so that $|y_k-y|<\varepsilon$ for $k \geqq K_2$, then for $k \geqq \max\{K_1, K_2\}$ it follows from the intermediate value theorem that f_k has at least two disjoint crossing intervals in I_m at y_k-level. This contradiction implies that E_{mn} is closed.

Next we show that E_{mn} is nowhere dense. Let $f \in C[0,1]$ and $\varepsilon>0$ be given. By the uniform continuity of f, there exists $\delta>0$ such that $|x'-x''|<\delta$ implies $|f(x')-f(x'')|<\varepsilon/4$. Numbers $x_0, x_1, \ldots, x_k$ may be chosen so that $x_0=0$, $x_k=1$ and $0<x_{i+1}-x_i<\delta$ for $i=0,1, \ldots, k-1$. We define a continuous piecewise linear function g by requiring that $g(x_i)=f(x_i)$ for $i=0,1, \ldots, k$ and that g be linear on $[x_i, x_{i+1}]$ for $i=0,1, \ldots, k-1$. It is easily verified that $\|f-g\|<\varepsilon/2$. If g is linear on $J=[a, b]$ and $\sup(g \mid J)-\inf(g \mid J)<\varepsilon/4$, then we choose x_1, x_2 and x_3 so that $a<x_1<x_2<x_3<b$. We define $h(a)=g(a), h(x_1)=h(x_3)=g(a)-\varepsilon/4$, $h(x_2)=g(b)+\varepsilon/4$ and $h(b)=g(b)$. We further require that h be linear on $[a, x_1]$, $[x_1, x_2]$, $[x_2, x_3]$ and $[x_3, b]$. It follows that $\nu(h, J, y) \leqq 1$ fails for $\inf(h \mid J)<y<\sup(h \mid J)$. Since I_m can be partitioned into closed intervals J having the properties above, we can construct a function h on I_m by requiring that $h \mid J$ be as defined above. If we let $h(x)=g(x)$ for $x \in I_m'$, then $\|g-h\|<\varepsilon/2$ and $h \in E_{mn}'$. Thus, E_{mn} is nowhere dense.

Let $E_{m0}=\{f: f$ is constant on $I_m\}$. It is straightforward to show that E_{m0} is closed and nowhere dense, and we omit the proof.

We next prove that $(\bigcup_{m=1}^{\infty} \bigcup_{n=0}^{\infty} E_{mn})' \subseteq B$. Suppose that $f \in(\bigcup_{m=1}^{\infty} \bigcup_{n=0}^{\infty} E_{mn})'$. Clearly, f is not constant on any interval by the way E_{m0} is defined. Let x_0 be chosen so that f has neither a local max nor a local min at x_0. Let I be an open interval containing x_0. Then there exist m and n such that x_0 is in the interior of I_m relative to $[0,1]$, $I_m \subseteq I$ and

$$\inf(f \mid I_m)+\frac{1}{n} \leqq f(x_0) \leqq \sup(f \mid I_m)-\frac{1}{n}.$$

Since $f \notin E_{mn}$, it follows that $f^{-1}(f(x_0)) \cap I_m$ contains at least two points. Thus f is locally recurrent at x_0. Let $F_{mn}=B' \cap E_{mn}$ for $m=1,2, \ldots$ and $n=0,1,2, \ldots$. It follows that $F=\bigcup_{m=1}^{\infty} \bigcup_{n=0}^{\infty} F_{mn}$ is of the first category and $B=F'$. This completes the proof.

We recall that $f \in C[0,1]$ has a strict local max (min) at x_0 in case there exists a deleted neighborhood N of x_0 such that $f(x_0)>f(x)$, $(f(x_0)<f(x))$ for $x \in N$.

The set of points at which $f \in C[0,1]$ has a strict local max or min is at most countable (see, for example, Saks [4, p. 261]). The latter fact will enable us to take $f \in B$ and modify it to obtain a function in $C[0,1]$ which is nonconstant and locally recurrent everywhere. Let M be the set of points at which f has either a strict local max or min. In case M is finite, a sketch of the proof proceeds as follows. We may imagine that the graph of f is modified by having a horizontal segment of length one inserted wherever a strict local max or min occurs. If the resulting function has its domain $[a, b]$ transferred to $[0,1]$ by composition with $g(x)=(b-a)x+a$, it will then have the desired properties. Next, we consider in detail the case in which M is countable. The proof is similar in spirit to the proof of the previous case, but it is more complex. We let $M=\{x_i\}_{i=1}^{\infty}$ and proceed by defining a mapping from $[0,1]$ to the collection whose members are either closed intervals or singletons in $[0,2]$. For each $x \in[0,1]$, let

$$r(x)=\begin{cases}[a_i, b_i] & \text{for } x=x_i \\ \left\{x+\sum_{j \in J(x)}\left(\frac{1}{2}\right)^j\right\} & \text{for } x \neq x_i\end{cases}$$

where $J(x)=\{j: x_j<x\}$, $a_i=x_i+\sum_{j \in J(x_i)}\left(\frac{1}{2}\right)^j$ and $b_i=a_i+\left(\frac{1}{2}\right)^i$.

LEMMA 1. *The mapping r has the following properties:*

(1) *It is strictly increasing in the sense that* $z_1 < z_2$ *implies* $y_1 < y_2$ *for all* $y_1 \in r(z_1)$, $y_2 \in r(z_2)$.

(2) *It is upper semicontinuous with respect to inclusion. That is, for every* $\varepsilon > 0$ *and* $x_0 \in [0,1]$, *there exists* $\delta > 0$ *such that* $|x - x_0| < \delta$ *and* $x \in [0,1]$ *implies* $r(x) \subseteq [\inf r(x_0) - \varepsilon, \sup r(x_0) + \varepsilon]$.

(3) $$\bigcup_{x\in[0,1]} r(x) = [0,2].$$

Proof: Part (1) follows directly from the definition of r. For the proof of part (2), let $\varepsilon > 0$ be given. Let $\delta < \varepsilon/2$ be chosen so that

$$\sum_{i=K}^{\infty} (\tfrac{1}{2})^i < \frac{\varepsilon}{2} \text{ where } K = \min\{j: 0 < |x_j - x_0| < \delta\}.$$

For $x_0 - \delta < x < x_0$ we have

$$\inf r(x_0) - \inf r(x) \leqq x_0 - x + \sum_{i=K}^{\infty} (\tfrac{1}{2})^i < \varepsilon$$

and for $x_0 < x < x_0 + \delta$

$$\sup r(x) - \sup r(x_0) \leqq x - x_0 + \sum_{i=K}^{\infty} (\tfrac{1}{2})^i < \varepsilon.$$

If $\{j: 0 < |x_j - x_0| < \delta\} = \emptyset$, then the sums can be omitted. This completes the proof of part (2). It is clear that $\inf r(0) = 0$ and $\sup r(1) = 2$. If $0 < y < 2$ and $y \notin \bigcup_{x\in[0,1]} r(x)$ then we let $E = \{x: \sup r(x) < y\}$ and $F = \{x: y < \inf r(x)\}$. By part (2), E and F are open in $[0,1]$. We have that $0 \in E$ and $1 \in F$. Since $E \cup F = [0,1]$ and $E \cap F = \emptyset$, we have contradicted the fact that $[0,1]$ is connected. This completes the proof of Lemma 1.

Next, we define a mapping $t: [0,2] \to [0,1]$ by requiring $t(y) = x$ if and only if $y \in r(x)$. Parts (1) and (3) of Lemma 1 imply that t is well defined on $[0,2]$. It follows directly that $t^{-1}(x) = r(x)$ for $x \in [0,1]$ and that $t(0) = 0$ and $t(2) = 1$.

LEMMA 2. *The mapping t has the following properties:*

(1) *For* $y_0 \in (\bigcup_{i=1}^{\infty} r(x_i))'$, $y_0 < y$ *implies* $t(y_0) < t(y)$ *and* $y < y_0$ *implies* $t(y) < t(y_0)$.

(2) *It is continuous.*

Proof: Suppose that $y_0 < y$. If $t(y) = t(y_0)$, then $r(t(y)) = r(t(y_0)) = \{y_0\}$ and we have a contradiction. If $t(y) < t(y_0)$, then we may apply Lemma 1, part (1), with $z_1 = t(y)$, $z_2 = t(y_0)$ to obtain $y < y_0$. This is a contradiction. A symmetric argument applies in case $y < y_0$. This completes the proof of part (1).

If t fails to be continuous at y_0, then there exists $\{y_i\}_{i=1}^{\infty}$, $y_i \to y_0$ as $i \to \infty$, and $\delta > 0$ such that $|t(y_i) - t(y_0)| \geqq \delta$ for $i = 1, 2, \ldots$. By taking a subsequence of $\{y_i\}_{i=1}^{\infty}$ if necessary, we may assume that there exists $x_0 \in [0,1]$ such that $t(y_i) \to x_0$ as $i \to \infty$. But $y_i \in r(t(y_i))$ for $i = 1, 2, \ldots$ and Lemma 1, part (2) imply that $y_0 \in r(x_0)$. That is, $t(y_0) = x_0$, and we have a contradiction. This completes the proof of Lemma 2.

Next, we show that $f \circ t: [0,2] \to [0,1]$ is continuous, nonconstant and everywhere locally recurrent. Suppose that $y_0 \in r(x_i)$ for some i. Then $f \circ t(y) = f \circ t(y_0)$ for all $y \in r(x_i)$ and $f \circ t$ is locally recurrent at y_0. Suppose next that

$$y_0 \in \left(\bigcup_{i=1}^{\infty} r(x_i)\right)'.$$

Let a and b be such that $a < y_0 < b$. By Lemma 2, part (1), $t(a) < t(y_0) < t(b)$. Since f is locally

recurrent at $t(y_0)$, there exists $z_1 \neq t(y_0)$, $t(a) < z_1 < t(b)$ such that $f(z_1) = f(t(y_0))$. The intermediate value theorem and Lemma 2, part (2), imply the existence of $y_1 \neq y_0$, $a < y_1 < b$, such that $t(y_1) = z_1$. Hence $f \circ t(y_1) = f \circ t(y_0)$. The usual modifications in case $y_0 = 0$ or $y_0 = 2$ complete the proof that $f \circ t$ is locally recurrent on $[0,2]$. Since f is not constant, there exist z_1 and z_2, $z_1 < z_2$, such that $f(z_1) \neq f(z_2)$. By Lemma 1, part (1), we have that $y_1 < y_2$ for any y_1, y_2 such that $y_1 \in r(z_1)$ and $y_2 \in r(z_2)$. Also, $f \circ t(y_1) \neq f \circ t(y_2)$. Lemma 2, part (2), implies that $f \circ t$ is continuous. If we let $g(x) = 2x$, $0 \leqq x \leqq 1$, then it follows directly that the function $f \circ t \circ g \in C[0,1]$ is nonconstant and everywhere locally recurrent. Thus, we have proved the following result.

THEOREM 3. *There exists a function in $C[0,1]$ which is nonconstant and everywhere locally recurrent.*

References

1. K. A. Bush, Locally recurrent functions, this MONTHLY, 69 (1962) 199–206.
2. ———, Pathological functions, crinkly and wild, J. Math. Anal. Appl., 35 (1971) 559–562.
3. E. Hewitt and K. Stromberg, Real and Abstract Analysis, Springer-Verlag, New York, 1965.
4. S. Saks, Theory of the Integral, Hafner, New York, 1937.

DEPARTMENT OF MATHEMATICS, EASTERN WASHINGTON STATE COLLEGE, CHENEY, WA 99004.

SPHERES WHICH LOOK ROUND ALL ROUND ARE ROUND

E. H. KRONHEIMER

1. Introduction. Let us agree — at the risk of appearing insensitive to the finer points of mathematical terminology — to call the surface bounding a ball in E^3 (= Euclidean 3-space) a **round sphere**, in order to distinguish it from a **sphere**, by which we will just mean a topological 2-sphere embedded in E^3: in other words, any subspace of E^3 homeomorphic to a round sphere. We shall make the same distinction in E^2 between **circles** (= simple closed curves) and **round circles**, and in E^4 between **hyperspheres** (= topological 3-spheres) and **round hyperspheres**; but **lines**, **planes**, and **hyperplanes** (= 3-planes) will all be flat as usual.

If it isn't round, a sphere can look very peculiar: that is to say, its image under orthogonal projection into a plane in E^3 may be very far from being a topological disc. This may of course be a symptom of a pathological embedding — as in the case of Alexander's Horned Sphere [**1**, 4–6] — but even a nicely embedded sphere, such as the surface of a loosely tangled piece of rope, can have an image of some topological complexity.

Nevertheless, the image always has an "outline" (we shall make that word precise in a moment) which is a circle. This circle is certainly a round circle if the sphere is a round sphere; but the converse also holds: no sphere, however weirdly embedded, has exclusively round outlines unless it itself is round.

In the course of proving these (not perhaps surprising) assertions, it will become clear how they can be extended — in particular to higher-dimensional spheres and spaces.

The argument will rest on two results: the Jordan-Brouwer Theorem (*Any n-sphere in E^{n+1}, where $n \geqq 1$, separates E^{n+1} into two components of each of which it is the frontier* [**3**, 4.7]), and the following classical lemma [**2**, 61.II.4]:

LEMMA. *Let Z be a compact, connected and locally connected subspace of the plane P which cannot be disconnected by removing a point. Then the frontier of each component of $P \backslash Z$ is a circle.*

2. Propositions. We begin by defining the word "outline."

Let X be a compact subset of E^n ($n \geqq 2$). Let P be a plane in E^n and π the orthogonal projection of E^n onto P. Since $\pi(X)$ is bounded, $P \setminus \pi(X)$ has exactly one unbounded component U. The **outline** of X in P is the frontier of $P \setminus U$ in the topological space P. Since $\pi(X)$ is closed, it contains this frontier.

PROPOSITION 1. *Let S be a sphere. Then every outline of S is a circle.*

Proof. Let π be the orthogonal projection of E^3 onto the plane P, and let $Z = \pi(S)$. Since $\pi \mid S$ is a closed map, Z is a quotient space of S and consequently inherits from S the properties of being compact, connected and locally connected. We can therefore deduce from the lemma that the outline of S in P is a circle, provided we verify that, whenever $z \in Z$, the set $Z \setminus \{z\}$ is connected.

We know from the Jordan-Brouwer Theorem that $E^3 \setminus S$ has two components. Call the bounded component T: then S is the frontier of T. Any line intersecting T must intersect S; so π maps T into Z. Put L for the line $\pi^{-1}(z)$: then $Z \setminus \{z\}$ is the image under π of $(S \cup T) \setminus L$, which is the closure in $E^3 \setminus L$ of $T \setminus L$. Hence $Z \setminus \{z\}$ is connected if $T \setminus L$ is connected — and that $T \setminus L$ is connected follows from the elementary result that no open subset of E^3 can be disconnected by removing a line.

PROPOSITION 2. *Let S be a sphere every outline of which is round. Then S is round.*

Proof. Given any plane P, the outline C of S in P is a round circle; consequently a plane Q orthogonal to P will touch C if and only if it supports S. ("Q touches C" means that $Q \cap P$ is tangent to C; "Q supports S" means that S intersects Q but does not intersect both components of $E^3 \setminus Q$.)

Now let P and P' be two planes, and let C and C' be the respective outlines of S in P and P'. Choose a plane Q orthogonal to both P and P', and let Q_1 and Q_2 be the two planes parallel to Q which touch C. Then Q_1 and Q_2 both support S: hence they also touch C'. This implies that C and C' have the same diameter $2r$ (= distance between Q_1 and Q_2); so r does not depend on the choice of P. The normals N and N' to P and P' through the centres of C and C' cannot be parallel unless they coincide; furthermore, they both lie in the plane Q_0 parallel to, and equidistant from, Q_1 and Q_2: hence they intersect. Since this is true for each choice of P and P', every normal constructed in this way will intersect every other such normal. Therefore, since the normals are not all coplanar, they all have a unique point in common: call this point a.

It now follows (i) that the outline of S in any plane through a is a round circle with centre a and radius r; and hence (ii) that S is entirely contained in the closed ball B^* in E^3 with centre a and radius r. Denote the round sphere which bounds B^* by S^*, and let $p \in S^*$. Choose a plane P^+ through p and a, and let π be the orthogonal projection of E^3 onto P^+. By (i) p belongs to the outline of S in P^+ and hence to $\pi(S)$. Since the only point of B^* which π maps to p is p, we deduce from (ii) that $p \in S$. Hence $S \supset S^*$.

That $S = S^*$ — and therefore that S is round — now follows from the observation — itself an immediate corollary of the Jordan-Brouwer Theorem — that no proper subspace of a sphere is a sphere.

3. Fewer outlines. We do not really need to know that *every* outline is round to prove Proposition 2. If L is a specified line in E^3, it is enough to postulate that the outline of S *in every plane containing* L should be round. (Thus, for example, a planet which rotates about an axis orthogonal to the line joining the planet to earth, and which always looks round, is actually round, and is not just a body whose cross-sections in planes parallel to the equator are arbitrary discs of constant width.)

For if we use the same argument as before to prove this new form of Proposition 2, we find that the only substantive effect of the weakened hypothesis is this: since all the normals corresponding to N and N' which can now be constructed are in fact coplanar, it is no longer clear that they must intersect in a unique point a. However, Q_1 is now the same plane for all admissible choices of P and P' ($\neq P$); and since the normals to C and C' in the plane Q_1 must both contain the non-empty set $S \cap Q_1$, it

follows that their point of intersection is independent of the choice of P and P' — which in turn implies that N and N' meet in a point a independent of this choice.

The rest of the proof remains valid, provided the planes considered in statement (i) — and the plane P^+ — are all required to contain the line through a parallel to L.

4. More objects. It is clear from the proof that in Proposition 1 we can replace the condition '*Let S be a sphere*' by '*Let S be a compact, connected and locally connected set which is the frontier in E^3 of a bounded, open and connected set T.*' These conditions on S and T cannot, however, be reduced any further. For example, if we take S to be the surface of a comb (of finite thickness) with countably many teeth $D_0, D_1, D_2, \ldots$, where tooth D_n is of height 1, of width 4^{-n}, and (for $n > 0$) at distance 2^{-n} from tooth D_0, then S has an outline which is not locally connected and therefore is not a circle.

In the proof of Proposition 2 it was not until the last paragraph that we needed to know anything about S beyond its property of being a compact set in E^3 with round circles as outlines. This property alone, therefore, is enough to show that $S = S^* \cup D$, where S^* is a round sphere and D consists of interior points of the ball B^* bounded by S^*. To ensure that $D = \varnothing$ it is sufficient to require S to be the outer surface in E^3 of some compact subset X of E^3. Here the **outer surface** of a compact subset of E^n $(n \geqq 3)$ in a hyperplane is defined precisely as we defined the outline of a compact subset of E^n $(n \geqq 2)$ in a plane: in particular, the outer surface of X in E^3 itself is the frontier of the unbounded component W of $E^3\backslash X$. Indeed, with this requirement on S, the component W is a connected and unbounded subset of $E^3\backslash S^*$ and hence has no points in B^*: so S has none in the interior of B^*, and therefore $S = S^*$.

The observation that the outlines of S must be identical with those of X leads to the following variant of the result we have just established:

Proposition $2^\dagger$. *Let X be a compact set in E^3 and suppose that every outline of X is a round circle. Then the outer surface of X is a round sphere.*

5. Other dimensions. In E^2 the obvious counterpart to Proposition 1 is trivially true, whereas the equally obvious counterpart to Proposition 2 is no less trivially false. (*Every* 0-sphere is round.) But this is quite different from the situation in higher dimensions, which is made clear by considering what happens in E^4.

Proposition 1 has the following E^4-version — which is false:

Conjecture 1a. *Let H be a hypersphere. Then every outer surface of H is a sphere.*

As a counterexample we can construct a hypersphere H whose image under a suitable projection is a solid torus, much as we would construct a sphere in E^3 to project onto an annulus. (Explicitly, take H to be the boundary of the topological 4-cube

$$\{(t+\theta,\ r\cos\theta,\ r\sin\theta,\ z)\mid (r-2)^2+z^2\leqq 1,\ 0\leqq\theta\leqq 2\pi,\ 0\leqq t\leqq 1\},$$

and project into the hyperplane $\{0\}\times\mathbf{R}^3$.)

However, Proposition 1 also has the following *valid* extension to E^4:

Proposition 1b. *Let H be a hypersphere. Then every outline of H is a circle.*

Proof. In the proof of Proposition 1 change "E^3" to "E^4", "S" to "H", and "line" to "plane".

Finally there are also two E^4-versions of Proposition 2. Both of these are true.

Proposition 2a. *Let H be a hypersphere every outer surface of which is a round sphere. Then H is round.*

Proof. Provided we change "E^3" to "E^4", "S" to "H", "outline" to "outer surface", "sphere" to "hypersphere", "circle" to "sphere", "plane" to "hyperplane", and "line" to "plane", the proof of Proposition 2 applies, except at one point in the second paragraph: we cannot deduce that the normals

N and N' intersect from the fact that they lie in the same hyperplane Q_0. To remedy this, we choose a second hyperplane R, not parallel to Q, which is orthogonal to both P and P'. Then N and N' lie in a hyperplane R_0 parallel to R as well as in the hyperplane Q_0 parallel to Q: hence they lie in the plane $Q_0 \cap R_0$ and intersect as before.

PROPOSITION 2b. *Let H be a hypersphere every outline of which is round. Then H is round.*

Proof. Let M be a hyperplane in E^4 and P be a plane in M. Denote the image of H under orthogonal projection into M by X. Then (i) H and X have the same outline in P, and (ii) H and X have the same outer surface in M. From (i) and Proposition 2† we deduce that the outer surface of X is a round sphere. Therefore, from (ii) and Proposition 2a, it follows that H is round.

References

1. J. G. Hocking and G. S. Young, Topology, Addison-Wesley, Reading, Mass., 1961.
2. K. Kuratowski, Topology, vol. II, Academic Press, New York, 1968.
3. E. H. Spanier, Algebraic Topology, McGraw-Hill, New York, 1966.

DEPARTMENT OF MATHEMATICS, BIRKBECK COLLEGE, LONDON WC1E 7HX, ENGLAND.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics and Statistics, The University of Calgary, Calgary, Alberta, Canada, T2N 1N4. (From July 1976 to June 1977: Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, England.)

A TOURNAMENT DESIGN PROBLEM

JENIFER HASELGROVE AND JOHN LEECH

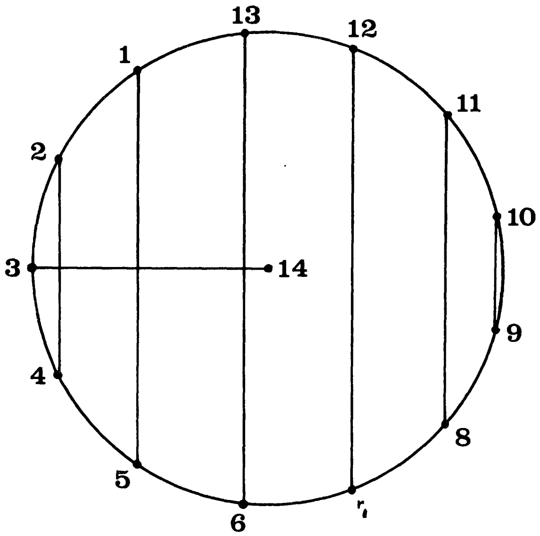
A round robin tournament is played among $2n$ players in $2n-1$ rounds. There are n courts of unequal attractiveness available for the matches, and each round is played at one time using all the courts.

PROBLEM: arrange the matches so that no player plays more than twice on any one court.

Gelling and Odeh [**2**] have discussed the case $n = 4$, and Robert Gray of Johnstone, Scotland (unpublished) has independently proposed the problem and solved it for $2n-1 \not\equiv 0 \pmod 3$.

We have first to arrange the matches so that they are played in $2n-1$ rounds of n matches each. A simple solution is given by Kraitchik [**3**, p. 227], Lockwood [**4, 5**], and Brooke [**1**], among others. We call this the **circle design** and illustrate it for the case $n = 7$. We number the players from 1 to 14 and draw a diagram as shown in the figure, with player 14 represented as the centre of a circle and players 1 through 13 represented as points evenly spaced on the circumference. A match between players i and j is represented by a line joining the points i and j. The matches to be played in round k are those

represented by the radius $(14, k)$ and the chords (i, j) which are perpendicular to it. The figure shows the matches for round 3. The whole tournament is given in Table 1, whose rows are the rounds. In the kth row the entry in column 0 corresponds to the radius $(14, k)$ and the subsequent entries to the chords in order, beginning with the one nearest to the point k.



We note first that the problem of the courts for an odd number of players is easily solved. For 13 players, for example, we remove player 14 from Table 1, so that column 0 represents the players not involved in the rounds, and then the other columns represent the courts in a solution.

0		1		2		3		4		5		6	
1	14	**2**	**13**	3	12	4	11	5	10	6	9	7	8
2	14	1	3	4	13	**5**	**12**	6	11	7	10	8	9
3	14	2	4	1	5	6	13	7	12	**8**	**11**	√9	10
4	14	3	5	2	6	1	7	8	13	9	12	**10**	**11**
5	14	4	6	3	7	2	8	**1**	**9**	10	13	11	12
6	14	5	7	**4**	**8**	3	9	2	10	1	11	12	13
7	14	6	8	5	9	4	10	3	11	2	12	1	13
8	14	7	9	**6**	**10**	5	11	4	12	3	13	1	2
9	14	8	10	7	11	6	12	**5**	**13**	1	4	2	3
10	14	9	11	8	12	7	13	1	6	2	5	**3**	**4**
11	14	10	12	9	13	1	8	2	7	**3**	**6**	4	5
12	14	11	13	1	10	**2**	**9**	3	8	4	7	5	6
13	14	**1**	**12**	2	11	3	10	4	9	5	8	6	7

TABLE 1

For the 14 players, columns 1 through 6 represent valid arrangements for courts, with each player except number 14 appearing exactly twice, but column 0 does not. Gray's solution is to exchange the entry marked by bold face type in each row with the entry in column 0 of the same row. The middle row has no exchange. The pattern of exchanges is clear from the table. The entry selected in row i, $i \neq 7$, contains the numbers $3i-1 \pmod{13}$ and $14-i$. Thus the selected entries contain between them

two of each of the numbers 1 through 13, except 7. When all the exchanges have been made, column 0 will have two of each number except for one of 14 and one of 7, and will therefore be a valid court arrangement. Each other column receives from column 0 two of the numbers it has lost and two of 14, and so it is also a valid court arrangement.

Gray's solution fails when $2n-1\equiv 0 \pmod 3$, for then the set of numbers $3i-1 \pmod{2n-1}$, instead of containing one each of 1 through $2n-1$, contains three each of $2,5,8,\ldots,2n-2$. The problem therefore remains of finding solutions for these cases.

Gray's solution is symmetric in an obvious sense. By taking only the circle design and restricting the search to symmetric solutions, we have found (without computational aids) solutions for $n=8,11,14$ and 17. This was done unsystematically; we have no general formula. For $n=5$ we have found both symmetric and unsymmetric solutions.

As well as finding single solutions, there is the problem of finding all essentially distinct solutions. For $n=2$ there is no solution. For $n=3$ the tournament design is essentially unique and so is the court solution.

W. D. Wallis (in Wallis, Street and Wallis [6], pp. 91–93) has proved that there are six distinct tournament designs for $n=4$. (A tournament design is equivalent to a 1-factorization of the complete graph on $2n$ vertices.) He has characterized the designs by the number d_3 of sets of three rounds which form two separate sub-tournaments of four players each, and the number d_2 of pairs of rounds which form part of such a pair of sub-tournaments but for which the third member of the set is absent. An example of each design is given in Table 2. Gelling and Odeh found that the designs 2(a) and 2(d) do not admit solutions to the court problem but the others do. Design 2(f) is the circle design; we have found that there are six distinct solutions, of which two are symmetric in the sense used above. We have enumerated the solutions for the other designs but we do not know how many are distinct.

1 2	3 4	5 6	7 8
1 3	2 4	5 7	6 8
1 4	2 3	5 8	6 7
1 5	2 6	3 7	4 8
1 6	2 5	3 8	4 7
1 7	2 8	3 5	4 6
1 8	2 7	3 6	4 5

$d_3=7,\quad d_2=0$

(a)

1 2	3 4	5 6	7 8
1 3	2 4	5 7	6 8
1 4	2 3	5 8	6 7
1 5	2 6	3 8	4 7
1 6	2 5	3 7	4 8
1 7	2 8	3 6	4 5
1 8	2 7	3 5	4 6

$d_3=3,\quad d_2=4$

(b)

1 2	3 4	5 6	7 8
1 3	2 4	5 7	6 8
1 4	2 3	5 8	6 7
1 5	2 7	3 6	4 8
1 6	2 8	3 5	4 7
1 7	2 6	3 8	4 5
1 8	2 5	3 7	4 6

$d_3=1,\quad d_2=4$

(c)

1 2	3 4	5 6	7 8
1 3	2 4	5 7	6 8
1 4	2 3	5 8	6 7
1 5	2 8	3 6	4 7
1 6	2 7	3 5	4 8
1 7	2 6	3 8	4 5
1 8	2 5	3 7	4 6

$d_3=1,\quad d_2=6$

(d)

1 2	3 4	5 6	7 8
1 3	2 4	5 7	6 8
1 4	2 5	3 8	6 7
1 5	2 7	3 6	4 8
1 6	2 8	3 7	4 5
1 7	2 3	4 6	5 8
1 8	2 6	3 5	4 7

$d_3=0,\quad d_2=3$

(e)

1 2	3 8	4 7	5 6
1 3	2 4	5 8	6 7
1 4	2 6	3 5	7 8
1 5	2 8	3 7	4 6
1 6	2 3	4 8	5 7
1 7	2 5	3 4	6 8
1 8	2 7	3 6	4 5

$d_3=0,\quad d_2=0$

(f)

TABLE 2

For $n = 5$ we have enumerated only the symmetric solutions for the circle design; there are 15 essentially distinct ones. Gelling and Odeh found that there are 396 distinct tournament designs; we have found isolated solutions for some of them.

The problem of enumerating distinct solutions for $n > 4$ seems intractably large, but it would be interesting to find a systematic method giving at least one solution for the cases $2n - 1 \equiv 0 \pmod 3$.

We are indebted to the referee for drawing our attention to the paper by Gelling and Odeh.

References

1. M. Brooke, Solution to a pairing problem, this MONTHLY, 74 (1967) 201.
2. E. N. Gelling and R. E. Odeh, On 1-factorizations of the complete graph and the relationship to round robin schedules, Proc. Third Manitoba Conference on Numerical Math., Winnipeg, (1973) 213–221.
3. M. Kraitchik, Mathematical Recreations, Allen and Unwin, London, 1943.
4. E. H. Lockwood, American tournament, Note 1213, Math. Gaz., 20 (1936) 333.
5. ———, Tournament problems, Note 3027, Math. Gaz., 46 (1962) 220.
6. W. D. Wallis, A. P. Street and J. S. Wallis, Combinatorics; Room Squares, Sum-Free Sets, Hadamard Matrices, Lecture Notes in Mathematics Vol. 292, Springer-Verlag 1972.

COMPUTING SCIENCE DEPARTMENT, UNIVERSITY OF GLASGOW, SCOTLAND.
COMPUTING SCIENCE DEPARTMENT, UNIVERSITY OF STIRLING, SCOTLAND.

CLASSROOM NOTES

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Laboratoire Calcul des Probabilités, Université de Paris, T.56, 4 Place Jussieu, 75–230 Paris, France.

DOUBLE INTEGRALS AS INITIAL VALUE PROBLEMS

JOHN M. KARON AND JAMES V. RALSTON

Traditionally we prove theorems concerning the definite integral *via* its definition as a limit of sums, and we motivate the definitions of area, volume, work, etc., as definite integrals in the same way. Alternatively, many of these definitions of physical and geometrical quantities can be motivated by an appropriate initial value problem (IVP) for a differential equation, an approach used by Levi **[1]** and Loomis **[2a, 2b]** for single integrals. In fact some of the theorems can also be proved by obtaining an appropriate IVP.

We wish to point out that this IVP method can also be used for double integrals. In particular, we shall give a rather simple proof of the change-of-variables formula, a proof which does not depend on ideas from advanced calculus. First we will review the method for the single integral and add several ideas to those presented by Loomis.

We emphasize that the difference lies in the spirit of the derivations, not the ingredients of the proofs. We consistently study definite integrals by looking at how they change as one limit of integration changes, à la Newton and Leibnitz, rather than using the method of exhaustion, à la Archimedes.

In each case, the function f to be integrated is assumed to be continuous.

1. Single integrals. Our basic approach is found in the usual intuitive proof of the fundamental theorem of integral calculus, using the interpretation of the definite integral as area. Suppose we wish

to find the area under the graph of $y = f(t)$ between $t = a$ and $t = b$, where $f(t) \geqq 0$. Let $A(x)$ be the area between $t = a$ and $t = x$. Then ΔA is approximately a strip of width Δx and height $f(x)$, so by letting $\Delta x \to 0$ we see that $A(x)$ is the solution of the intial value problem $A'(x) = f(x)$, $A(a) = 0$. Hence, if F is any antiderivative of f, then $A(x) = F(x) - F(a)$, and this is equal to the proper definite integral by the fundamental theorem.

It is just as easy to compute volumes. To find the volume obtained by rotating this area about the x-axis, let $V(x)$ be the volume obtained by rotating $A(x)$. Then $V(x + \Delta x) - V(x)$ is the volume of a 'slice' of thickness Δx, so

$$V'(x) = \pi f(x)^2, \qquad V(a) = 0.$$

If we rotate the area about the y-axis, ΔV is a 'shell', and

$$V'(x) = 2\pi x f(x), \qquad V(a) = 0.$$

This approach is nothing more than a precise version of the usual differential formulation ($dA = f(x)dx$, etc.). However, it emphasizes a very nice analogy: the derivative of the area function is the length of the moving boundary; the derivative of the volume function is the area of the moving boundary (for both 'slicing' and 'shells'). One can take the analogous approach to work, force on the side of a dam, arc length, and surface area.

We can also prove the change-of-variable formula. Let $g(u)$ be continuously differentiable and let

$$I(x) = \int_{g(a)}^{g(x)} f(t)dt$$

(the limits have been chosen this way to prepare for the transformation $t = g(u)$). Then $I'(x) = f(g(x))g'(x)$ and $I(a) = 0$, hence

$$\int_{g(a)}^{g(x)} f(t)dt = \int_a^x f(g(u))g'(u)du.$$

2. Double integrals. To understand the following proofs, a student need know only the basic facts about double integrals: a definition for integrals over rectangles (perhaps as iterated single integrals) and Fubini's theorem for the change of order of integration for such integrals, some definition of the double integral over an arbitrary region, and the mean value theorem for a double integral. As an aside, we note that Fubini's theorem for a rectangle can be proved by the method used here for volume and the change-of-variables formula.

To prepare for our proof of the change-of-variables formula, we first motivate the double integral form for the volume between the surface $z = f(s, t) \geqq 0$ and the s-t plane. We show this for a region with a rectangular base; it can be extended to an arbitrary region by the method of exhaustion (as in the limit-of-sums approach). Let $R(x, y)$ be the rectangle in the s-t plane with opposite vertices at the origin and (x, y). Let $V(x, y)$ be the volume of the solid with base $R(x, y)$ and top, the surface $z = f(s, t)$. We consider

$$\{V(x + h, y + h) - V(x + h, y) - V(x, y + h) + V(x, y)\}/h^2 \tag{1}$$

which approaches $V_{xy}(x, y)$ as $h \to 0$ (we will show that this partial derivative exists). The numerator in (1) is the volume of a solid with base an h-by-h square with one (fixed) vertex (x, y) and opposite vertex $(x + h, y + h)$. This volume lies between $h^2 \min f$ and $h^2 \max f$, where the minimum and maximum are taken over the square. By the intermediate value theorem, the volume is $h^2 f(P)$, where P is some point in the square. Letting $h \to 0$, we obtain the initial value problem

$$\frac{\partial^2 V}{\partial x \partial y} = f(x, y), \qquad V(x, 0) = V(0, y) = 0. \tag{2}$$

Clearly the solution is

$$V(x,y)=\int_0^x ds\int_0^y f(s,t)dt. \tag{3}$$

Another approach, which also shows that V is differentiable, begins with

$$\{V(x+h,y)-V(x,y)\}/h. \tag{4}$$

The numerator is the volume of a solid with base a strip of width h and height y. Let $g(w,y)$ be the cross-sectional area cut by the plane $s=w$:

$$g(w,y)=\int_0^y f(w,t)dt.$$

Then the numerator of (4) is bounded below by $h\ \min g(w,y)$ and above by $h\ \max g(w,y)$; the minimum and maximum are taken for $x\leqq w\leqq x+h$. Letting $h\to 0$, we see that the partial of V with respect to x exists and

$$\frac{\partial V}{\partial x}=\int_0^y f(x,t)dt.$$

This also yields (2).

3. The change-of-variables formula. Our first derivation of this formula uses the same idea as the first derivation of the volume formula, but we apply (1) to the function

$$F(x,y)=\iint_{R^*(x,y)} g(u,v)du\,dv,$$

where $R^*(x,y)$ is the image of $R(x,y)$ under the transformation $u=u(s,t)$, $v=v(s,t)$ and $g(u,v)=f(u(s,t),v(s,t))$. We shall assume that the transformation is globally one-to-one, that both u and v have two continuous derivatives, and that the Jacobian $J(x,y)$ of this transformation does not vanish. The quotient corresponding to (1) is now

$$\frac{1}{h^2}\iint_{R^*(h)} g(u,v)du\,dv; \tag{5}$$

$R^*(h)$ is the image of the h-by-h square with opposite vertices (x,y) and $(x+h,y+h)$. By using first-order Taylor expansions with remainder for u and v, we shall show in the next paragraph that $R^*(h)$ is almost the parallelogram determined by the vectors

$$(hu_x+O(h^2),hv_x+O(h^2)) \quad\text{and}\quad (hu_y+O(h^2),hv_y+O(h^2))$$

from the point $(u(x,y),v(x,y))$; the partial derivatives are also evaluated at (x,y). The area of this parallelogram is the length of the cross-product of these vectors, which is $|h^2J(x,y)+O(h^3)|$. Therefore, the integral of $g(u,v)$ over this parallelogram is $h^2 f(P)|J(x,y)|+O(h^3)$, where P is a point in the parallelogram. The limit of (5) as $h\to 0$ will yield

$$\frac{\partial^2 F}{\partial x\partial y}=f(x,y)|J(x,y)|, \qquad F(x,0)=F(0,y)=0, \tag{6}$$

provided we can show that $R^*(h)$ is close enough to the parallelogram. The solution to (6) is then analogous to that for the volume problem.

To obtain the details concerning the parallelogram, we use the Taylor expansion with remainder

$$u(s,t)=u(x,y)+u_x(x,y)(s-x)+u_y(x,y)(t-y)+O(h^2) \qquad |x-s|,|y-t|\leqq h$$

and the analogous expansion for v. The images of the vertices of the h-by-h square under the Taylor expansions form the vertices of the parallelogram. The Taylor expansions also show that the images of the sides of the square are within $O(h^2)$ of the corresponding sides of the parallelogram. Therefore, each curve bounding $R^*(h)$ lies within a strip of length $O(h)$ and width $O(h^2)$ centered on the corresponding side of the parallelogram, so the area of $R^*(h)$ is in fact $h^2|J(x,y)|+O(h^3)$, and (6) is valid.

The following proof of the change-of-variables formula shows that the second mixed partial of F exists and is continuous, but it is aesthetically less pleasing, as we need the definition of the single definite integral as a limit of sums. Here is an outline. We use

$$\frac{F(x+h,y)-F(x,y)}{h}=\frac{1}{h}\sum_{i=1}^{n}\iint\limits_{R_i^*(h)} g(u,v)du\,dv \qquad (n=y/h),$$

where $R_i^*(h)$ is the image of the h-by-h square bounded by the lines $s=x$, $s=x+h$, $t=(i-1)h$, and $t=ih$. Using the Taylor series as above, the ith term in the sum is

$$f(P_i)h^2|J(x,(i-1)h)|+O(h^3),$$

where P_i is a point in the corresponding parallelogram. As $h\to 0$, the sum approaches a definite integral and the quotient on the left approaches $F_x(x,y)$, so

$$\frac{\partial F}{\partial x}=\int_0^y f(x,t)|J(x,t)|\,dt.$$

4. Concluding remarks. Some comments are in order on the relation of our proof of the change-of-variables formula for double integrals to other proofs.

Most proofs make essential use of the fact that the image of a very small rectangle under the transformation to (u,v) coordinates is approximately a parallelogram, either as we have, or in deriving the integral representation (in terms of the Jacobian) for the area of the image of an arbitrary region under the transformation. (Osserman **[3]** obtains the latter *via* Green's theorem.) Our proof of this fact, using Taylor series, is equivalent to representing the transformation locally by its differential. For an outline of, and references to, various proofs of the change-of-variables formula, see Schwartz **[4]**.

Some of the more traditional proofs can be used to show that, if f is integrable over the original domain, then it is integrable over the image. Our proof depends on f being continuous, and does not show integrability.

We reiterate that it is our point of view which is different, not our ingredients. Of course, the same proof could be used for the change-of-variables formula in higher dimensions, but the numerator of (1) would become increasingly complicated.

References

1. H. Levi, Polynomials, Power Series, and Calculus, Van Nostrand, Princeton, 1968.
2a. L. Loomis, Calculus, Addison-Wesley, Reading, Mass., 1974.
2b. ———, Introduction to Calculus, Addison-Wesley, Reading, Mass., 1975.
3. R. Osserman, Two-Dimensional Calculus, Harcourt, Brace & World, New York, 1968.
4. J. Schwartz, The formula for change in variables in a multiple integral, this MONTHLY, 61 (1954) 81–85.

DEPARTMENT OF MATHEMATICS, COLORADO COLLEGE, COLORADO SPRINGS, CO 80903.
DEPARTMENT OF MATHEMATICS, UCLA, LOS ANGELES, CA 90024.

ON A NECESSARY AND SUFFICIENT CONDITION FOR RIEMANN INTEGRABILITY

LEO M. LEVINE

Let D be the set of points of discontinuity of $f: I \to R$ where I is any interval of the real line R. Let $L = \{x: f \text{ has a left-hand limit at } x\}$.

La Vita [1] proved that if S is a closed subset of L then $D \cap S$ is countable. He then made use of this result to prove that f is continuous a.e. if and only if f has a left-hand limit a.e. Thus Lebesgue's well-known necessary and sufficient condition for Riemann integrability of a bounded function on a finite interval can be replaced by the "weaker" condition that f have a left-hand limit a.e.

It is the purpose of this note to present a much shorter and simpler proof than that given in [1] of the following stronger result, from which the equivalence of the two conditions for Riemann integrability follows immediately.

THEOREM. *$D \cap L$ is countable.*

Proof. Let $D_n = \{x: \operatorname{osc}(f, x) > 1/n\}$, $n = 1, 2, 3, \ldots$, where

$$\operatorname{osc}(f, a) = \lim_{\delta \to 0+} (\sup\{f(x): |x - a| < \delta\} - \inf\{f(x): |x - a| < \delta\}).$$

Since $D = \bigcup_{n=1}^{\infty} D_n$, and therefore $D \cap L = \bigcup_{n=1}^{\infty} D_n \cap L$, we need only prove that $D_n \cap L$ is countable for each n.

Suppose $x_0 \in D_n \cap L$. Since $x_0 \in L$, there exists a $\delta > 0$ such that $|f(x) - f(x_0^-)| < 1/2n$ for all $x \in (x_0 - \delta, x_0)$. Hence

$$|f(x_1) - f(x_2)| < 1/n \quad \text{for} \quad x_1, x_2 \in (x_0 - \delta, x_0).$$

It follows that if $x \in (x_0 - \delta, x_0)$ then $\operatorname{osc}(f, x) \leqq 1/n$, so that $x \notin D_n$. Thus any point of $D_n \cap L$ is the right endpoint of an open interval which contains no point of $D_n \cap L$. Since these open intervals are clearly disjoint, and hence form a countable set, it follows that $D_n \cap L$ is countable, and the theorem is proved.

References

1. J. A. La Vita, A necessary and sufficient condition for Riemann integration, this MONTHLY, 71 (1964) 193–196.

DEPARTMENT OF MATHEMATICS, QUEENSBOROUGH COMMUNITY COLLEGE, BAYSIDE, NY 11364.

A BOUNDED DERIVATIVE WHICH IS NOT RIEMANN INTEGRABLE

CASPER GOFFMAN

We give a very simple example of a bounded derivative which is not Riemann integrable. Let $f: [0, 1] \to R$ be defined as follows. Let $G \subset [0, 1]$ be a dense open set which is the union of pairwise disjoint open intervals, $\{I_n\}$, the sum of whose lengths is $\frac{1}{2}$. For each n, let $J_n \subset I_n$ be a closed interval in the center of I_n such that the lengths satisfy $l(J_n) = [l(I_n)]^2$. For each n, define f on J_n to be continuous, 1 at the center, 0 at the end points and always between 0 and 1. Define f to be 0 everywhere else.

The function f is not Riemann integrable. For if π is any partitioning of $[0, 1]$ the intervals of π in

which the oscillation of f is 1 have length sum exceeding $\frac{1}{2}$, so that

$$\overline{\int} f - \underline{\int} f \geqq \frac{1}{2}.$$

That f is a derivative follows from the fact that it is bounded and approximately continuous so that it is the derivative of its indefinite Lebesgue integral $F(x) = \int_0^x f(t)dt$. It is of interest that this may be shown without the use of the Lebesgue integral. The function F may be obtained as an improper integral by letting

$$F(x) = \sum_{n=1}^{\infty} \int_{K_n} f(t)dt,$$

where $K_n = J_n \cap [0, x]$, $n = 1, 2, \ldots$. Let $I \subset [0, 1]$ be an interval which meets the complement of G, and let n be such that $I \cap J_n \neq \emptyset$. Let $S_n = l(I_n)$. Since $S_n \leqq \frac{1}{2}$, it follows that $l(I \cap I_n) \geqq \frac{1}{2}(S_n - S_n^2) \geqq \frac{1}{4}S_n$. Then $l(I \cap J_n) \leqq l(J_n) = S_n^2 \leqq 16\{l(I \cap I_n\}^2$. If $N = \{n: I \cap J_n \neq \emptyset\}$, then

$$\sum_{n \in N} l(I \cap J_n) \leqq \sum_{n \in N} 16\{l(I \cap I_n)\}^2 \leqq 16\{l(I)\}^2.$$

For $x \notin G$ and $y \neq x$, we have $\int_x^y f(t)dt \leqq 16(y - x)^2$, whence $F'(x) = 0$. That $F'(x) = f(x)$ on G is obvious.

For comparison see **[1]**, **[2]**, and **[3]**.

References

1. C. Goffman, Real Functions, Rinehart, New York, 1960, p. 210–11.

2. E. W. Hobson, The Theory of Functions of a Real Variable and the Theory of Fourier Series, Vol. 2, Cambridge, p. 412–421.

3. I. P. Natanson, Theory of Functions of a Real Variable, vol. 1, Ungar, New York, 1960, p. 133.

DEPARTMENT OF MATHEMATICS, PURDUE UNIVERSITY, WEST LAFAYETTE, IN 47907.

HOMOLOGICAL DOTS

STEVEN H. WEINTRAUB

In this note, I would like to introduce a new mathematical game. I originally called it "Commutative Diagrams and Exact Sequences," a name which accurately reflects its essence, but for reasons of euphony, as well as its resemblance to the children's game of Dots, I have changed the name to the one above.

The rules are as follows:

1. The game is played on a 5×5 lattice.

2. The first move consists of writing down an arbitrary finitely generated abelian group at some lattice point.

3. Any subsequent move consists of writing down an arbitrary finitely generated abelian group on a vacant lattice point which is orthogonally adjacent to an already occupied lattice point, together with maps to/from the adjacent group(s), subject to the following conditions:

(a) All arrows are to the right or down.

(b) All horizontal or vertical sequences must be exact.

(c) All squares must commute.

4. Scoring: Suppose a player makes a move which makes it impossible to legally fill some other

lattice point. Then that player scores that lattice point and moves again. (Otherwise, players move alternately.) The one who scores the most points wins.

To illustrate, here are the first few moves of a sample game.

1. Player 1 plays $\mathbf{Z}$ at $(3,3)$.
2. Player 2 plays $\mathbf{Z}\oplus\mathbf{Z}$ at $(3,4)$ and $f_1\colon \mathbf{Z}\to\mathbf{Z}\oplus\mathbf{Z}$ by $f_1(x)=(2x,3x)$.
3. Player 1 plays $\mathbf{Z}_5$ at $(4,3)$ and $f_2\colon \mathbf{Z}\to\mathbf{Z}_5$ by $f_2(x)=x \bmod 5$.
4. Player 2 can now score, and does! He plays $\mathbf{Z}$ at $(2,4)$ and $f_3\colon \mathbf{Z}\to\mathbf{Z}\oplus\mathbf{Z}$ by $f_3(x)=(x,0)$. He scores the point $(4,4)$. The board looks as in Figure I, where player 2 has indicated his scoring $(4,4)$ by writing his name thereupon.

$$
\begin{array}{cccccc}
1 & \cdot & \cdot & \cdot & \cdot & \cdot \\
2 & \cdot & \cdot & \cdot & \mathbf{Z} & \cdot \\
 & & & & \downarrow f_3 & \cdot \\
3 & \cdot & \cdot & \mathbf{Z} \xrightarrow{f_1} & \mathbf{Z}\oplus\mathbf{Z} & \cdot \\
 & & & \downarrow f_2 & & \\
4 & \cdot & \cdot & \mathbf{Z}_5 & 2 & \cdot \\
5 & \cdot & \cdot & \cdot & \cdot & \cdot \\
 & 1 & 2 & 3 & 4 & 5
\end{array}
$$

FIG. I

Proof of scoring: Suppose we were to write a group G at $(4,4)$, with maps $g_1\colon \mathbf{Z}\oplus\mathbf{Z}\to G$ and $g_2\colon \mathbf{Z}_5\to G$. By exactness, $\ker(g_1)=\{(x,0)\in\mathbf{Z}\oplus\mathbf{Z}\}$. Now $\ker(g_1)\cap\operatorname{im}(f_1)=\{(0,0)\}$, so g_1f_1 is an injection, $g_1f_1\colon\mathbf{Z}\to G$. But by commutativity, $g_1f_1=g_2f_2$. But f_2 is not an injection, so this is impossible.

Obviously, there are many possible variants of this game. For example, one may change the board size, or instead of scoring as above play that the first person unable to make a legal move loses, or allow arbitrary abelian groups (or for that matter, play the game in the abelian category of your choice). In practice, the game as described above has proved to be an interesting one.

DEPARTMENT OF MATHEMATICS, LOUISIANA STATE UNIVERSITY, BATON ROUGE, LA 70803.

MATHEMATICAL EDUCATION

EDITED BY PAUL T. MIELKE AND SHIRLEY HILL

Material for this Department should be sent to Paul T. Mielke, Department of Mathematics, Wabash College, Crawfordsville, IN 47933, or to Shirley Hill, Department of Mathematics, University of Missouri, Kansas City, MO 64110.

APPLIED MATHEMATICS: AN INTRODUCTION VIA MODELS

RONALD F. BARNES

Introduction. In light of the recent job crunch for mathematically trained people, especially in the teaching field, students in mathematics are beginning to express much more interest in applications of mathematics to the physical and social sciences. Certain of the social sciences including economics and

QUASI-MODULAR APPROACH IN MATHEMATICS FOR THE DISADVANTAGED*

G. H. GAONKAR, BEVERLY DOUGLAS AND K. KRISHNAN

1. Introduction. For that segment of the student body which has received substandard high school training, the conventional mathematics instruction in colleges is inadequate, inaccessible, and often impersonal. Commitments by several universities to meeting the needs of their service areas and to open-enrollment policies have significantly increased the enrollment of these students called "the disadvantaged." However, the integrity of the higher education process warrants a learning system which will equip these students with the fundamentals of precollege mathematics so that they can accommodate to the rigors of science training. Accordingly, a learning system has been designed at Southern Illinois University at Edwardsville, to help the disadvantaged pursue careers in the sciences and engineering.

During past years, several academic institutions have started programs to increase the opportunities for disadvantaged students, programs involving special monetary awards, counseling, and tutorial services. In general such programs have met with varying degrees of success in reaching their goal — "preparing disadvantaged students for survival in the mainstream of American higher education," [1]. Tangible results include increased enrollment, improved grade point averages, and job opportunities. However, there is one facet to these programs which has recently received increased attention: how to prepare and motivate these students to study science areas such as engineering, computer and laboratory science, etc., where there is an ever-increasing demand, particularly for minority persons [2]. In this respect, these programs have been, in large measure, only marginally successful. Available statistics from science and engineering programs for disadvantaged students are indeed disappointing.

This is a disquieting situation which warrants more than a token solution, something more than "empty fronts to appease the outside pressures and to attract government funds" [3]. The number of disadvantaged students continues to rise on university campuses. This rise is significant at a time when the total enrollment, particularly at the graduate level, is on the decline. Disadvantaged "students are in the universities, and will be in the universities,...nor will the luxury of mathematics departments ignoring the needs of these students continue to exist" [3]. There are several reasons for this situation. One can mention a few to provide an insight into the inherent complexities in remediation programs. First, such programs are invariably classified as "special" or "supportive" and are often "founded in a climate of hostility with very little consideration given to making them a legitimate part of the academic mainstream" [4]. Second, in a university system where an instructor's prestige is measured by his publication and by his association with graduate programs and other research work, many instructors believe that remedial programs are neither challenging nor deserving of academic stature. Third, a student with a high school diploma likes to be treated as a college freshman or sophomore, and not as a disadvantaged student. Monotonous and drill-type exercises in high school arithmetic and algebra just add to the "demeaning and deleterious nature of the standard approach to mathematical remediation" [5]. It is also cited that colleges often "unwittingly place unnecessary bureaucratic obstacles in the path of a student so that he becomes confused, discouraged, or completely unable to cope with the system" [3].

The question is, how can we prepare and motivate them in mathematics under these circumstances? One approach is the time honored one — conventional classroom lectures with outside tutorial facilities. Bittinger [6] feels that this approach should not be overlooked as an effective means

* This is a condensed and expository version of a full report (quoted here as reference 12) without statistical subtleties. The full report containing extensive short-term statistics pertaining to disadvantaged students and students from the minorities is available upon request as Preprint No. 33, Box 65, SIU-E, Edwardsville, Illinois 62026.

of teaching remedial math. "The humanizing effect of an instructor who can relate to such students can help to change attitudes and bring success. Many remedial students respond well to a lecture approach because they need to 'see the instructor do it' " [6]. In spite of recent advances in the teaching of remedial courses by self-study approaches, our experience also shows that the conventional lecturing approach has some merit in reaching disadvantaged students. However, we must recognize that the initial preparation of these students varies rather drastically. There still is much research needed in synthesizing a learning system which, as a feedback control system, is based on the conventional lecturing method and which, at the same time, is responsive to the heterogeneous initial preparation of these students.

Thus, a "learning-system" has been devised which utilizes feedback from past experiences of failures and successes. Its design is very much influenced by earlier research efforts of Bittinger [6], Williams [3], Bohigian [5], and others. The emphasis is not remediation, nor is it based on the notion that these students are slow learners. Based on experience, it verifies that "what these students lack in preparation is more than compensated for by desire to overcome their deficiencies, a willingness to tackle new topics, and a rich potential that hitherto had hardly been tapped" [5].

2. Programmed instruction. One of the approaches in mathematical remediation is through programmed books, where tutorial services and other "self-learning kits" like video-tapes, audio visual techniques, etc., usually serve as necessary adjuncts. Experience has shown that these students are also ill-prepared in English; they can hardly differentiate between clauses and phrases, not to speak of reading mathematics books [7]. It is seldom possible to motivate a student if he cannot live up to the intellectual thresholds of his classroom, to its challenges, to its thought-provoking aspects. In this regard tutorial facilities and the so-called self-learning approaches through programmed books are of secondary importance, no matter how extensive. Not being initiated into the art of reading mathematics books, most of these students are invariably frustrated. "We already knew that our students never read mathematics textbooks but we discovered why: they are unreadable...The language and syntax are so complex that you have to understand the context being discussed before you can read the discussion of the concept. For beginning students and particularly previously unsuccessful students such a textbook can be fatal" [8]. According to Bittinger, "The further a remedial course moves from lectures to self-study, the greater the demands for verbal skills" [6]. The result is the path of least resistance — drop out. A few who are better motivated than the rest barely pass the course, even with the help of tutorial services. An unfortunate consequence concerns general studies and other basic courses in science and engineering. Once again, due to their inadequate background in preparatory mathematics and English, they are subject to a frustrating experience in basic courses which are often conducted on a large, impersonal basis.

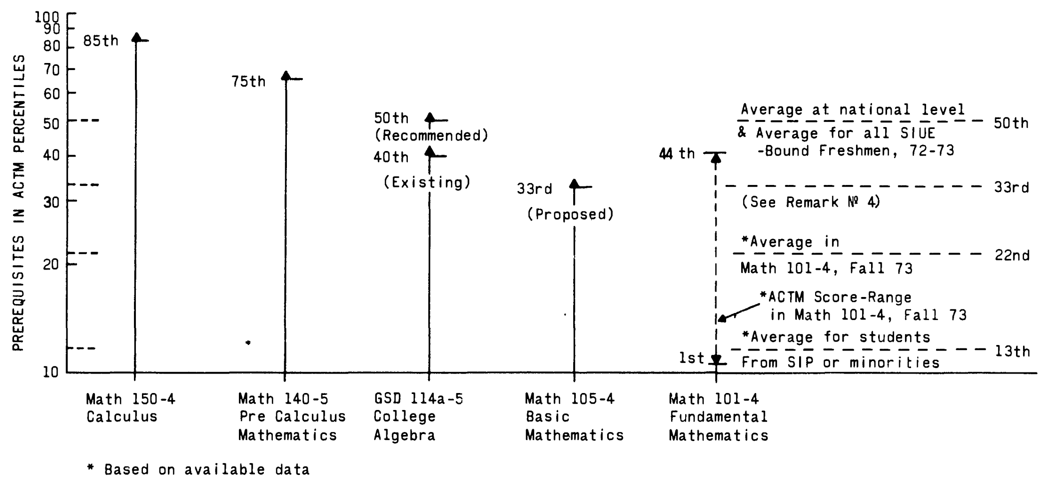
3. Quasi-modular approach (QMA). The learning system referred to earlier is called QMA. It involves a coordination of counselling and conventional classroom lectures with built-in workshop sessions. Tutorial facilities are available on request, and help-sessions are arranged for specially identified groups in response to their performances on weekly tests. When compared to several remediation efforts, it is comprehensive in several respects:

(1) Two specially designed remediation courses in mathematics are phased into the standard curriculum of the School of Science and Technology. The instructional process provides full and routine access to counselling and participation in workshop sessions. Appraisal of student-evaluation data indicates that this approach significantly alleviates the negative connotations associated with remediation programs.

(2) Deficiencies in arithmetical skills and elementary algebra are approached through specially prepared units or modules which "can be taught with some degree of sophistication" [5]. ("We mention in passing that many self-learning kits were found to be so elementary that the students were 'turned-off'.") The intent is not to give the student drill-type exercises, but rather to introduce basic concepts through project-type problems which "have some research potential appropriate for

students" [5]. Emphasis is placed upon the applications of arithmetic and algebra to applied problems from different disciplines. Some of these modules involve only rudimentary concepts of statistics.

(3) The duration of remediation is flexible and is tailored to individual needs. As shown in Figure 1, two courses are offered in remedial mathematics: (i) Fundamental Mathematics, starting with basic concepts in arithmetic, and dealing with elementary Algebra; and (ii) Basic Mathematics, starting with intermediate algebra and providing a natural transition to College Algebra which is a General Studies course in the science area. In combination with Precalculus Mathematics this sequence allows a student to take four quarters of Mathematics preparatory to Calculus. On the basis of two diagnostic tests and counselor-report evaluation, these students are grouped into four categories: (i) those students who require sequential Fundamental Mathematics and Basic Mathematics in two quarters; (ii) those who could take both math courses in one quarter on an accelerated basis; (iii) those who need only Basic Math; and (iv) those who could take College Algebra directly with regular assistance from help-sessions.



Remarks:

1. No prerequisites for Math 101–4. Attendance at help sessions during the first three weeks may be required.
2. In Math 101–4 and Math 105–4, the number of contact periods is 25% per credit higher than in the other three courses.
3. Math 105–4 was introduced from winter '74 on only.
4. Out of 1279 entering freshmen, 450 scored in the lower one-third range of the ACTM percentile range.

FIG. 1

(4) Remediation in English is a concomitant but independent part of this program. It is axiomatically accepted that the student must have writing skills and reading speed satisfactory to the English department before he can perform rewardingly in *any* academic program.

(5) An integral part of this program is the Science-Engineering Club which is designed to provide an orientation towards careers in science and engineering.

4. Format and structure of QMA. There exists an extensive literature on different approaches in mathematical remediation [6]: We mention in passing that the term "modular approach" in the literature does not seem to imply a well-defined and unique instructional process. Different types of modular approaches have been proposed by several investigators [6, 8, 9]. They do exhibit a considerable degree of freedom; for example see the Instructional Modular approach of Riner [9] and

the modular approach of Ablon **[8]**, etc. In the present paper the term "module" is used to denote a unit on a particular topic or on a particular section of a chapter etc., e.g., module on fractions or on fractions and decimals, etc. The construction of these modules is based on three criteria: (1) The examples should be challenging and not routine and drill-type in that they often relate to problems of applied interest. (2) These modules are to be used in the context of supplementary materials to alleviate the problem of heterogeneous backgrounds of students. They should help the students to read the prescribed textbooks which the instructor follows closely during lectures. (The students discuss the contents of these modules only during workshop sessions which are an integral part of the entire instructional process.) (3) The emphasis is on the readability and not so much on mathematical precision. (This criterion is equally applied in the selection of textbooks which contain applications to several disciplines.) The tacit assumption is that the students making use of these modules and the textbook are self confident neither in mathematics nor in English.

To elaborate further on the preparation of the modules, examples of two typical modules are given, one in arithmetic and one in intermediate algebra. The module on fractions and percentages refers to a telephone survey on supersonic transport aircraft. Students will be asked to compute the average favorable and unfavorable responses to the practicability of using such aircraft as long-range transport vehicles. The module contains the formula for computing averages with rudimentary explanations on the concept of averages. A wide range of such examples can be taken from elementary books on statistics such as *Mathematics Through Statistics*, by Auslander, *et al.* (The Williams and Wilkins Company, Baltimore, 1973.) The second example taken from reference 5 consists of expressing a given number using the integer 3 only. For example: $43 = 33 + 3/0.3$ and $48 = 3^3 + \Sigma_{(3+3)}$ where $\Sigma_n = 1 + 2 + 3 + \cdots + n$. Another module in algebra was taken from *Intermediate Algebra* by M. Lial and C. Miller (Scott Foresman & Co., 1972), chapter 12 entitled "Applications of Algebra." This example refers to the differential growth rate or the law of simple allometry in cancer. This example, bypassing certain biological and mathematical subtleties, is no more involved than the examples of the type $y = b2^x$. It is our belief that students should and could be exposed to such examples with elementary exposure to exponential and logarithmic functions.

5. Mode of instruction according to QMA. As stated earlier, the first course in precollege algebra is Fundamental Mathematics. In general, the attrition rate in this course used to be very high. This is not surprising, since these students, through no fault of their own, have received a training in high school mathematics which is just not adequate for a college program. Therefore, in Fundamental Math only minimal preparation in high school mathematics is assumed, say, not more than a familiarity with simple fractions. Students are given two diagnostic tests to assess their actual preparation in mathematics. During the first three weeks, intensive tutoring and help sessions are arranged for those who fared poorly in the diagnostic tests. Some of these make-up or help sessions are conducted by using the specially prepared modules referred to earlier. Results are encouraging in that some of the students from help sessions have caught up with the rest of the Fundamental Math class. In order to motivate the students in mathematics the applications of arithmetic and algebra to social, biological, and physical sciences are included. Another approach to alleviate the problem of heterogeneous background is to restrict students to designated sections of Fundamental Math depending upon their scores in ACTM (American College Testing in Mathematics) or diagnostic tests, etc. This approach is found to be not practical since the student has to take a total of 4 to 5 other courses according to a restricted schedule.

The remedial courses are conducted by one instructor and two assistants. This instructor is responsible for the lecture and assists in the coordination of module-distribution and in the arrangement of help sessions. After about 30 minutes of regular lecture in a class of about 40 to 45 students, the students are asked to solve certain examples in the class and are encouraged to discuss problems individually with the instructor or with one of the assistants.

6. Final remarks. Appraisal of a program of this type is often based on information for which there exists very little comparative data so that terms such as substantial improvement, degree of success, etc., have varying connotations **[10, 11]**. Further, it is generally accepted that the procedures for assessing and implementing the criteria for appraising remedial programs are complex **[10, 11]**. The variety of proposed procedures, though numerous, varies drastically due to the subjectivity that is inherent in the selection of such criteria **[10, 11]**. There is still much research needed to present these procedures in a unified perspective in that different procedures are compared with respect to one set of comparable statistics. Meanwhile, it does not seem justified to burden this expository paper with appraisals according to different criteria, although such appraisals are of importance in conjunction with short-term and long-term statistics. As noted earlier, in a full report (available upon request) detailed short-term statistics are presented with respect to three populations — students taking the first course in College Mathematics, disadvantaged students, and students from the minorities. The data pertains to pre-college preparation, attrition and grade distributions, and responses to follow-up mathematics courses for a sequence of three courses — College Algebra preceded by two courses in preparatory mathematics. With the information thus far available, the intent is to study comparative initial trends concerning attrition rate, student's overall performance in remedial and regular mathematics courses, grade distribution, and responses to follow-up mathematics courses. Though quantitatively the results will require further substantiation through long-term statistics, the established trends should remain valid.

Appraisal of short-term statistics indicates that disadvantaged students can be motivated in mathematics, not by conventional methods of instruction, but according to QMA as outlined in sections 3, 4, and 5. Early experience with the program indicates "some" success and feedback in achieving the program objectives:

(1) to provide a conceptual and skill base for further studies in science and engineering; (2) to help disadvantaged students pursue careers in science and engineering; (3) to continue to develop instructional materials, including those tailored to individual needs. The preparation of these materials will take into consideration the feedback from course evaluation by students and the observed discrepancy between the projected and the obtained results in students' performance in each course.

Acknowledgements. This research was supported by the office of Emil Jason, Assistant Vice President for Special Programs and Minority Affairs. Comments from Emil Jason, R. N. Pendergrass, Norval Wallace are gratefully acknowledged. We are also grateful to the referee, who suggested ways to make these ideas accessible to a wider readership without recourse to tables and data.

References

1. A. B. Grobman, Shortchanging the disadvantaged student, Science, 177(4046) (July 1972) 1.

2. T. M. Jenkins and R. S. Larsen, A minorities educational program for technologists in a research laboratory, IEEE Transactions on Education, E–16(2) (May 1973) 115.

3. E. R. Williams, Mathematics for the disadvantaged, this MONTHLY, 81, (June-July 1974) 648–659.

4. Emil Jason, The Cutting Edge, The first National Conference on Special Emerging Programs in Higher Education, November, 1973, p. 2.

5. H. E. Bohigian, Mathematical Remediation with Dignity, Intern. J. Math. Educ. in Science and Technology, 4 (1973) 403–412.

6. M. L. Bittinger, A comparison of approaches for teaching remedial courses at the college level, Math. Teacher (May 1972) 455–456.

7. R. B. Kane, The readability of mathematics text books revisited, Math. Teacher, 63 (November 1970) 579–581.

8. L. B. Ablon, A modular approach to preparatory mathematics, this MONTHLY, 79 (December 1972) 1126–1131.

9. John Riner, Individualizing mathematics instruction, this MONTHLY, 79 (January 1972) 77–86.
10. Larry A. Braskamp and Robert D. Brown, Evaluation of programs for blacks, Educational Record (Winter 1972) 51–58.
11. James Wilson, Standard tests very often measure the wrong things, Math. Teacher (April 1973) 295–370.
12. G. H. Gaonkar, Beverly Macon Douglas, Kuppanna Krishnan, Preliminary results from a mathematics program for the disadvantaged, SIU-E Papers in Mathematics and Mathematical Sciences, Preprint No. 33, May 19, 1975.

SUPPLEMENTAL INSTRUCTIONAL PROGRAM (SIP) AND DEPARTMENT OF MATHEMATICAL STUDIES, SOUTHERN ILLINOIS UNIVERSITY, EDWARDSVILLE, IL. 62026.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before June 30, 1977.

E 2641. *Proposed by Philip Straffin, Beloit College, Wisconsin*

Given a convex polygon, and a point p inside it, define $D(p)$ to be the sum of perpendicular distance from p to the sides of the polygon (extended if necessary). Characterize those convex polygons for which $D(p)$ is independent of p.

E 2642. *Proposed by Antonio Rocha, Belo Horizonte, Minas Gerais, Brazil*

Let x, y, z be integers such that

$$x^2 + y^2 = z^{2m}, \qquad (x, y) = 1,$$

where m is a positive integer. If $4m - 1 = p$ is a prime, show that $p \mid xy$.

E 2643. *Proposed by Harry D. Ruderman, Hunter College Campus School*

Show that for no integer $n > 1$, $2^n - 1$ divides $3^n - 1$.

E 2644. *Proposed by Solomon W. Golomb and Lloyd R. Welch, University of Southern California*

Let $A_n = ru^n + sv^n$ $(n \geqq 0)$ where r, s, u, v are integers, $u \neq \pm v$ and let P_n be the set of prime divisors of A_n. Show that the union P of all P_n is infinite.

E 2645. *Proposed by Jerrold W. Grossman, Oakland University*

A deck of N cards is shuffled according to the following scheme. The cards, labeled 1 through N, are placed in order in a row. Independent random integers $r_1, \ldots, r_N$ are chosen successively, $1 \leqq r_i \leqq N$, and after the choice of each r_i the card then in position i is interchanged with the card then in position r_i. What is the probability that card s ends up in position t after the shuffle is complete?

•E 2646. *Proposed by William Wernick, City College, New York*

Let $A_1, \ldots, A_n$ be vertices of a regular n-gon inscribed in a circle with center O. Let B be a point on arc A_1A_n and $\theta = \sphericalangle A_nOB$. If a_k is the length of the chord BA_k express $\Sigma_{k=1}^{n}(-1)^k a_k$ as a function of θ.

SOLUTIONS OF ELEMENTARY PROBLEMS

Prime Triplets

E 2561 [1975, 936]. *Proposed by Joel M. Simon, Philadelphia, Pennsylvania*

Let (p_1, p_2, p_3) be a prime triplet spaced by the common interval d. Show that if d is not a multiple of 6, then $p_1 = 3$ and necessarily the triplet is unique. Discuss the situation if d is a multiple of 6.

Solution by the Temple University Problem Solving Group. Obviously $p_1 \neq 2$ and so d must be even. Therefore, if $d \not\equiv 0 \pmod 6$ then $d \not\equiv 0 \pmod 3$. Since one of the integers $p_1, p_1 + d, p_1 + 2d$ must be divisible by 3 it follows that $p_1 = 3$. Thus for a fixed d such that $d \not\equiv 0 \pmod 6$ there can be at most one prime triplet p_1, $p_2 = p_1 + d$, $p_3 = p_1 + 2d$.

When $d \equiv 0 \pmod 6$ many prime triplets with "difference" d are known. It is conjectured (see D. Hensley and I. Richards, *On the compatibility of two conjectures concerning primes*, AMS Symposia in Pure Math. 24 (1973), 123–127 and A. Schinzel, *Remarks on the paper* ..., Acta Arith. 7 (1961), 1–8) that for every $d \equiv 0 \pmod 6$ there are infinitely many prime triplets (p_1, p_2, p_3) with $p_2 - p_1 = p_3 - p_2 = d$.

Also solved by Anders Bager (Denmark), Allen Beadle, Wayne Boucher, Paul Bruckman, Fred Buckley, Keith Burns (Australia), Michael Ecker, Thomas Elsner, Lorraine Foster, Donald Fuller, George Garrison, William Gorman, III, M. G. Greening (Australia), Frederick Humburg, M. S. Klamkin, Carolyn MacDonald, Robert McCarthy, Arthur Marshall, Louise Moser, T. E. Moore, Bob Prielipp, Ken Rebman, Nan-Shan Shou (Hong Kong), K. R. P. Singh (India), A. M. Vaidya (India), G. Williams, Jr., Aleksandras Zujus, and the proposer.

Editor's Comment. Both M. S. Klamkin and B. Prielipp refer to L. E. Dickson, *History of the Theory of Numbers*, vol. I, p. 425, where the answer to the first part of the problem is given. Several solvers refer to W. Sierpinski, *Elementary Theory of Numbers*, where it is shown that there exist infinitely many prime triplets. As noted by Klamkin, Sierpinski also gives the following examples of 10 and 13 primes in arithmetic progression:

$$199 + 210k, \qquad k = 0, 1, 2, \ldots, 9;$$

$$4943 + 60060k, \qquad k = 0, 1, 2, \ldots, 12.$$

Directed Monochromatic Paths

E 2562 [1975, 937]. *Proposed by N. C. K. Phillips, University of the Witwatersrand, South Africa*

Each of the $\binom{m}{2}$ edges of the complete graph on m vertices is assigned a direction and one of n colors in such a way that there is no monochromatic directed path $\overrightarrow{AB}$, $\overrightarrow{BC}$ of length 2. How large can m be in terms of n?

Solution and generalization by Seth Chaiken, graduate student, Massachusetts Institute of Technology. We shall solve a more general problem. Let $f(n)$ be the largest m such that the edges of the complete graph K_m can be directed and colored with n colors so that there is no monochromatic directed path of length k. Then we claim that $f(n) = k^n$.

First we show that $f(n) \geqq k^n$. Let $m = k^n$ and let $0, 1, \ldots, m-1$ be the vertices of K_m. We direct the edge $\{i, j\}$ by $i \to j$ if $i < j$. For each vertex i of K_m let $i(s)$ $(0 \leqq s \leqq n-1, 0 \leqq i(s) \leqq k-1)$ be the digits of i expressed in base k, i.e.

$$i = \sum_{s=0}^{n-1} i(s)k^s.$$

We color the edge $\{i, j\}$ of K_m by the rth color $(r = 1, 2, \ldots, n)$ if $i(s) = j(s)$ for $s \geqq r$ and $i(r-1) \neq j(r-1)$. If $i_0, i_1, \ldots, i_t$ is a monochromatic directed path (of color r) of K_m then

$$0 \leqq i_0(r-1) < i_1(r-1) < \cdots < i_t(r-1) \leqq k-1$$

and consequently $t \leqq k-1$. Hence, K_m has no directed monochromatic paths of length k.

Next we show that $f(n) \leqq k^n$ by induction on n. This is valid if $n = 0$. Suppose that $f(n) \leqq k^n$ and that for some $m > k^{n+1}$ the edges of K_m can be directed and colored with $n+1$ colors so that there is no monochromatic directed path of length k. We fix one color, say green, and define the subsets $S_0, S_1, \ldots, S_{k-1}$ of the vertex set of K_m as follows: $i \in S_t$ if and only if the longest directed green path ending in i has length t. At least one S_t has more than k^n elements. That S_t determines a complete subgraph of K_m which can be colored by n colors so that there are no monochromatic paths of length k. This contradicts our induction hypothesis.

Also solved by D. M. Bloom, Marianne Gardner, Landy Godbold, O. P. Lossers (Netherlands), St. Olaf Problem Group, and the proposer. A partial solution by Charles Chouteau.

Volume and Surface Area of a Solid

E 2563 [1975, 937]. *Proposed by J. Th. Korowine, Athens, Greece*

Let f_1 and f_2 be non-negative periodic functions of period 2π and let $h > 0$. Let $P_1(\theta)$ and $P_2(\theta)$ be the points whose cylindrical coordinates are $(f_1(\theta), \theta, 0)$ and $(f_2(\theta), \theta, h)$ respectively. Find integrals for the volume and surface area of the solid bounded by the planes $z = 0$, $z = h$, and the lines $P_1(\theta)P_2(\theta)$.

Solution by M. S. Klamkin, University of Waterloo, Ontario, Canada. If a point (r, θ, z) lies on the surface generated by the motion of the segment $P_1(\theta)P_2(\theta)$ then

$$r = r(z, \theta) = f_1(\theta) + \frac{z}{h}(f_2(\theta) - f_1(\theta)), \qquad 0 \leqq z \leqq h. \tag{1}$$

Then the volume we want to find is given by

$$V = \frac{1}{2}\int_0^{2\pi}\left(\int_0^h r(z, \theta)^2 dz\right) d\theta = \frac{h}{6}\int_0^{2\pi} (f_1^2 + f_1 f_2 + f_2^2) d\theta.$$

This can also be written in the form

$$V=\frac{h}{6}(B_1+4M+B_2),$$

where B_i $(i=1,2)$ are the areas of the bases and M is the area of the mid-cross section. Explicitly

$$B_i=\frac{1}{2}\int_0^{2\pi} f_i^2 d\theta \quad (i=1,2), \qquad M=\frac{1}{2}\int_0^{2\pi}\left(\frac{f_1+f_2}{2}\right)^2 d\theta.$$

The lateral surface area S is given by the well-known area integral in cylindrical coordinates

(2) $$S=\int_0^{2\pi}\int_0^h \sqrt{r^2+(rr_z)^2+r_\theta^2}\,dz\,d\theta,$$

where r_z and r_θ are the partial derivatives of $r=r(z,\theta)$. In our case we have

(3) $$r_z=\frac{1}{h}(f_2-f_1), \qquad r_\theta=f_1'+\frac{z}{h}(f_2'-f_1').$$

Then the entire area of the solid is B_1+B_2+S.

Also solved by L. Kuipers (Switzerland), and the proposer. Partial solutions (first part only) by Thomas Elsner, William Gorman, and Kurt Rusnak.

Editor's Comment. The expression $r^2+(rr_z)^2+r_\theta^2$ is a quadratic in z because of formulas (1) and (3). The formula for S which emerges after this integration is given by Kuipers but it is too complicated to state it here.

Covering of Four-valent Graphs

E 2564 [1975, 1009]. *Submitted by Daniel J. Kleitman, Massachusetts Institute of Technology*

Can one cover the vertices of any regular graph of degree 4 (every vertex in it has degree four) by disjoint arcs and stars? (A star consists of a vertex and all the arcs containing it; an arc covers both its ends.) [This problem is due to R. L. Graham and is a variation of a problem of Claude Berge.]

Solution by S. C. Locke, graduate student, University of Waterloo, Ontario. Let v and e be the number of vertices and the number of edges, respectively, of a regular graph of degree 4. If the graph can be covered by disjoint arcs and stars then we must have

$$(v,e)=m(1,4)+n(2,1)$$

for suitable non-negative integers m and n. For the complete graph K_5 we have $v=5$, $e=10$ and the system of equations $m+2n=5$, $4m+n=10$ has no integral solutions. Thus K_5 is not coverable.

Another counterexample with $v=23$ was submitted by Kleitman.

Editor's comment. For a regular graph G of degree 4 we have $e=2v$ and the solution of the system $m+2n=v$, $4m+n=2v$ is $m=3v/7$, $n=2v/7$. Thus the condition $7\mid v$ is necessary.

There are examples which show that this condition is not sufficient. Let (x,y) be an arc joining two vertices x and y of G. If this arc is used in a covering of G then the other 6 edges incident with x or y must be included in stars. Since these stars are disjoint it follows that (x,y) is not contained in a 4-circuit. But it is easy to construct a regular graph of degree 4 on 7 vertices such that every arc lies on a 4-circuit.

Regularizing a Bipartite Graph

E 2565 [1975, 1009]. *Submitted by Daniel J. Kleitman, Massachusetts Institute of Technology*

Given a bipartite graph on n and $2n$ vertices that is regular on either set (of degree $2k$ and k respectively), can one necessarily find n vertices of the second kind such that upon their removal along with the arcs containing them the remaining graph is regular of degree k? [This problem is due to T. Nemetz.]

Solution by H. W. Lenstra and O. P. Lossers, Eindhoven University of Technology, the Netherlands. The answer is negative. Indeed, let $n = m(2m-1)$, $k = 2m-1$ where m is an integer $\geqq 2$. For each natural number r let $S_r = \{1, 2, \ldots, r\}$. The first group of n vertices consists of the 2-element subsets $\{i, j\}$ of S_{2m}. The second group of $2n$ vertices consists of the elements (u, v) of the Cartesian product $S_{2m-1} \times S_{2m}$. The vertices $\{i, j\}$ and (u, v) are joined by an arc if and only if $v \in \{i, j\}$. This defines a bipartite graph G such that each vertex $\{i, j\}$ has degree $2k$ and each vertex (u, v) has degree k. Let G' be a graph obtained by deleting from G n vertices of the second group and all the incident arcs. Assume that G' is regular of degree k. Let a, b, c be the numbers of vertices of G' of the form $(u, 1)$, $(u, 2)$, $(u, 3)$, respectively. Since $\{1, 2\}$ has degree k in G' we must have $a + b = k$. Similarly, $a + b = b + c = c + a = k$. This is impossible since $k = 2m - 1$ is odd.

Also solved by the proposer.

Obtuse Pythagorean Triplets

E 2566 [1975, 1010]. *Proposed by Edvard Kramer, Ljubljana, Yugoslavia*

A triplet (a, b, c) of natural numbers is an *obtuse Pythagorean triplet* if a, b, c are the sides of a triangle ABC with $\measuredangle C = 120°$. Such a triplet is *primitive* if a, b, c have no common factor other than unity.

(i) Show that each positive integer except 1, 2, 4 and 8 can appear as the smallest member of an obtuse Pythagorean triplet.

(ii)* What positive integers can appear in primitive obtuse Pythagorean triplets?

A composite of solutions by David P. Robbins, Phillips Exeter Academy, and Kenneth L. Yocom, South Dakota State University. A triplet (a, b, c) of natural numbers is an obtuse Pythagorean triplet (OPT) if and only if

$$a^2 + ab + b^2 = c^2. \tag{1}$$

(ii) We claim that a positive integer occurs in a primitive OPT if and only if it is either odd and $\geqq 3$ or a multiple of 8.

The "if" part is immediate since

$$(2n+1,\ 3n^2+2n,\ 3n^2+3n+1), \tag{2}$$

$$(8n,\ 12n^2-4n-1,\ 12n^2+1) \tag{3}$$

are primitive OPT's for $n \geqq 1$.

If an even integer x occurs in a primitive OPT (a, b, c) then since c is odd we can assume that $x = a$ and that b is odd. The equation (1) gives

$$a(a+b) = c^2 - b^2 \equiv 0 \pmod 8$$

and hence 8 divides a.

It remains to show that 1 cannot occur in an OPT. If it did then the equation

$$b^2 + b + 1 = c^2$$

would have a solution in positive integers. This is not the case since the equation implies that $b^2 < c^2 < (b+1)^2$.

(i) Since $2n + 1 < 3n^2 + 2n < 3n^2 + 3n + 1$ for $n \geqq 1$ and $8n < 12n^2 - 4n - 1 < 12n^2 + 1$ for $n \geqq 2$ and the triples (2) and (3) are OPT's it suffices to prove that 8 does not appear as the smallest member of an OPT (a, b, c). Let us assume that $a = 8$. Then (1) becomes

$$48 = (c + b + 4)(c - b - 4).$$

But this equation has no solutions in positive integers such that $c > b \geqq 8$. This completes the proof.

Remark by Alan Wayne, New Port Richey, Florida. A triplet (a, b, c) of natural numbers is an acute Pythagorean triplet (APT) if a, b, c are the sides of a triangle ABC with $\measuredangle C = 60°$ and $a \neq b$. If (a, b, c) is an OPT then $(a, c, a+b)$ is an APT and all APT's can be obtained in this manner. Using this it is easy to show that the assertion (i) holds also for APT's. We have also that every positive integer except 1, 2 and 4 appears as the smallest member of an APT.

Also solved by Anders Bager (Denmark), Leon Bernstein, George Berzsenyi, David Bienenfeld (Israel), M. T. Bird, D. M. Bloom, Ezra Brown, A. Charnow, Miltiades Demos, Kay Dundas, Thomas Elsner, Jessie Ann Engle, M. G. Greening (Australia), L. Kuipers (Switzerland), V. Linis (Canada), S. C. Locke (Canada), William Markel, Marilyn McIntosh, William McKay, Dewey Moore, Ram Murty & Kumar Murty, Chester Palmer, David Penney, Problem Solving Group Bern (Switzerland), Richard Pulskamp, Samson Rosenzweig, K. R. P. Singh (India), E. Trost (Switzerland), University of South Alabama Problem Group, Alan Wayne, and Charles Wexler.

Partially solved by Eugen Peter Bauhoff (Germany), Walter Bluger (Canada), G. A. Heuer, O. P. Lossers (Netherlands), Bob Prielipp, David Wright, and the proposer.

Editor's comment. Several respondents referred to L. E. Dickson, *History of the Theory of Numbers* (Vol. II, pp. 404–407) where the solutions of the diophantine equation $a^2 + ab + b^2 = c^2$ are discussed.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Hill Center, Rutgers University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate signed sheets and should be mailed before June 30, 1977.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

6138. *Proposed by Harry D. Ruderman, Hunter College Campus School*

Let $p_1, p_2, \ldots$ be consecutive primes with $p_1 = 2$. Then

(1) for every n there is a k for which $\Pi_{i=n}^{n+k} p_i$ is an abundant number.

(2)* Find an upper bound for k in terms of n.

6139. *Proposed by D. P. Munro, University of Sydney, Australia*

Consider a first-order predicate calculus, and all the relational structures appropriate to that calculus.

(a) Let $P_1, \ldots, P_k$ be a finite collection of mutually exclusive and exhaustive axiomatizable properties (so every relational structure has exactly one of the properties P_i). Must any of the P_i be in fact finitely axiomatizable, and if so, how many?

(b) As for (a), but with a countably infinite collection of mutually exclusive and exhaustive axiomatizable properties.

6140*. *Proposed by F.S. Cater, Portland State College, Oregon*

Let f be a continuous real-valued function on $[0,1]$ and let E_f denote the (possibly void) set $\{x \in [0,1]: f'(x) \text{ exists and is finite}\}$. Let $a(f) =$ Lebesgue outer measure of $f([0,1]\backslash E_f)$,

$$m(t) = \begin{cases} f'(t) & \text{for } t \in E_f \\ 0 & \text{otherwise.} \end{cases}$$

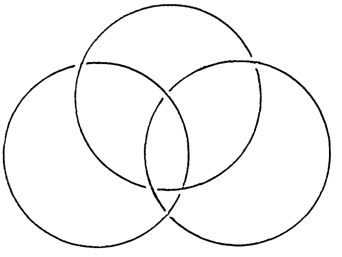
Let $b(f) = a(f) + \int_0^1 m(t)dt$ and $c(f) = a(f) + \int_0^1 [1 + m(t)^2]^{1/2} dt$.

Find (1) max $c(f)$ over all f such that $b(f) = 1$, and (2) min $c(f)$ over all f such that $b(f) = 1$. Describe functions for which $c(f)$ takes one of these values.

6141. *Proposed by Dennis Johnson and Herbert Taylor, Jet Propulsion Laboratory*

The Borromean Rings, shown in the figure, cannot be drawn without crossing on a surface of genus

1 (the torus), but they can on a surface of genus 3 (the three hole torus). Can the Borromean Rings be drawn without crossing on a surface of genus 2?



6142. *Proposed by L. O. Chung, North Carolina State University*

Find a function $f:[0,1]\to[0,1]$ which is continuous everywhere except on two countable dense subsets D_1, D_2 of rationals such that on D_1, f is right continuous but not left continuous, and on $D_2 f$, is left continuous but not right continuous.

6143. *Proposed by A. L. Macdonald, Eastern Michigan University*

The familiar method of fair division of a pie by passing a knife over it until someone is satisfied suggests the problem: Let $\pi_1, \pi_2, \ldots, \pi_n$ be non-atomic probability measures on a set X. Then there are pairwise disjoint sets $B_1, B_2, \ldots, B_n$ with $\pi_i(B_i) \geqq 1/n$.

SOLUTIONS OF ADVANCED PROBLEMS

Distributive Lattices

6032 [1975, 529]. *Proposed by D. J. Johnson, Air Force Institute of Technology*

Given that L and M are distributive lattices and $[\mathcal{G}, \leqq]$ is the partially ordered set of lattice morphisms from L to M ordered according to the rule $f \leqq g$ if and only if, for all x in L, $f(x) \leqq g(x)$ in M (i.e., $f(x) \wedge g(x) = f(x)$), is $[\mathcal{G}, \leqq]$ necessarily a lattice?

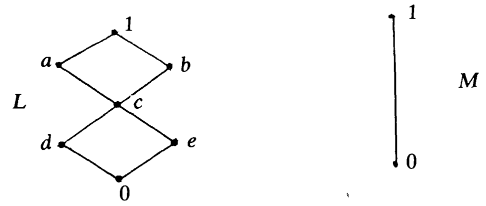
I. *Solution by F. David Hammer, Sherman Oaks, California.* The answer is no, and a familiar lattice supplies the reason. Let N be the distributive lattice with universe the set of natural numbers and operations g.c.d. and l.c.m. (designated $\wedge$ and $\vee$). Each permutation of the set of primes induces an automorphism of N as follows:

If f is such a permutation, extend it to N by setting

$$\bar{f}(p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_n^{\alpha_n}) = (f(p_1)^{\alpha_1})(f(p_2)^{\alpha_2})\cdots(f(p_n)^{\alpha_n}).$$

Let the primes be partitioned somehow into three element subsets. Let f be a permutation of the primes which is of order three on each member of the partition. We show that $\bar{f}$ and $\bar{f}^{-1}$ have no least upper bound, in fact no upper bound at all. Suppose on the contrary, that $\bar{f}, \bar{f}^{-1} \leqq g$ for some endomorphism g of N. Let p, q and r be primes, all larger than $g(1)$ such that $\{p, q, r\}$ is a member of the partition. Then $\bar{f}(p)$ and $\bar{f}^{-1}(p)$ both divide $g(p)$, so rq divides $g(p)$, similarly rp divides $g(q)$, hence r divides $g(p) \wedge g(q) = g(p \wedge q) = g(1)$, which contradicts the choice of r. Thus $[G, \leqq]$ is not even a directed partial order.

II. *Solution by A. A. Jagers, Technische Hogeschool Twente, Enschede, Netherlands.* Let L and M be the distributive lattices of the diagram below. Put $fx = 1$ if $x = 1$, a, resp. $gy = 1$ if $y = 1$, b, and $fx = gy = 0$ otherwise. Then $f, g \in \mathcal{G}$ and $\{f, g\}$ has two incomparable minimal upper bounds p and q in $\mathcal{G}$, viz. $px = 0$ if $x = e$, 0, resp. $qy = 0$ if $y = d$, 0, and $px = qy = 1$ otherwise.



Also solved by Sydney Bulman-Fleming (Canada), Thomas Kucera & Bill Sands (Canada), and the proposer.

Editor's note. Both the proposer and Susan Montgomery note that the affirmative statement of the problem appears as an exercise in Birkhoff and Bartee, *Modern Applied Algebra*, p. 273, no. 9.

Condition for $|f(z)|<1, |z|<1$

6033 [1975, 529]. *Proposed by Sanford Miller, State University of New York, Brockport*

Let $w(z)$ be regular in the unit disk D with $w(0)=0$ and let A be a complex number such that $\operatorname{Re} A \geqq 1$. If $z \in D$, show that

$$|w^2(z)+Aw(z)+zw'(z)|<1 \quad \text{implies} \quad |w(z)|<1.$$

Solution by Robert Vermes, McGill University. We can assume a little less, namely $\operatorname{Re} A = a \geqq \frac{1}{2}(\sqrt{5}-1)$. Let $g(z)=w^2(z)+Aw(z)+zw'(z)$, then $g(0)=0$ and $|g(z)|<1$ in $|z|<1$. Hence by the lemma of Schwarz, $|g(z)| \leqq |z|$ in $|z| \leqq 1$. From

$$\lim_{z\to 0}\left|\frac{g(z)}{z}\right| = |(A+1)w'(0)| \leqq 1$$

we obtain

$$|w'(0)| = \frac{1}{|A+1|} \leqq \frac{1}{a+1} < 1$$

which implies that there exists a maximal r_0, $0<r_0\leqq 1$ such that $|w(z)|<|z|$ for $0<|z|<r_0$. If $r_0=1$, we have $|w(z)|<1$ in $|z|<1$. The assumption $r_0<1$ leads to a contradiction. If $r_0<1$, then for some point z_0 with $|z_0|=r_0$, $|w(z_0)|=|z_0|$; without loss of generality we may assume that $z_0=r_0$. Multiply $g(r)$ by

$$r^{A-1} = \operatorname{Exp}[(a-1)\log r]\operatorname{Exp}(i \operatorname{Im} A \log r),$$

then

$$|r^{A-1}g(r)| = \left|r^{A-1}w^2(r)+\frac{d}{dr}[r^A w(r)]\right| \leqq |r^{A-1}|\,|r| = r^a$$

for $0<\varepsilon\leqq r\leqq r_0$. But

$$\left|\lim_{\varepsilon\to 0}\int_\varepsilon^{r_0} r^{A-1}g(r)dr\right| = \left|\lim_{\varepsilon\to 0}\int_\varepsilon^{r_0} r^{A-1}w^2(r)dr + r_0^A w(r_0)\right|$$

$$\leqq \lim_{\varepsilon\to 0}\int_\varepsilon^{r_0} |r^{A-1}g(r)|dr \leqq \lim_{\varepsilon\to 0}\int_\varepsilon^{r_0} r^a dr = \frac{r_0^{a+1}}{a+1}.$$

Hence

$$|r_0^A w(r_0)| \leqq \frac{r_0^{a+1}}{a+1} + \lim_{\varepsilon\to 0}\int_\varepsilon^{r_0} |r^{A-1}w^2(r)|dr \leqq \frac{r_0^{a+1}}{a+1} + \lim_{\varepsilon\to 0}\int_\varepsilon^{r_0} r^{a+1}dr = \frac{r_0^{a+1}}{a+1} + \frac{r_0^{a+2}}{a+2}.$$

Now $|w(r_0)| = r_0 < 1$, hence

$$1 \leqq \frac{1}{a+1} + \frac{r_0}{a+2} < \frac{1}{a+1} + \frac{1}{a+2} \leqq 1$$

if $a \geqq \frac{1}{2}(\sqrt{5}-1)$, a contradiction.

Also solved by S. K. Bajpar (India), G. H. Fricke, L. E. Mattics, Malcolm Robertson, St. Olaf Problem Solving Group, and the proposer.

Editorial notes. (1) Mattics, having obtained the better condition on A given in the solution above, raises the question as to the best possible constant A for which the conclusion is valid. (2) Robertson proves the following proposition: Let $w(z)$ be analytic in the unit disk D with $|w(0)| < 1$. Let A_z and B_z be complex numbers, possibly dependent on z, subject to the restraint:

If $B_z = 0$ let $|A_z| \geqq 2$; if $B_z \neq 0$ let $|A_z + B_z| \geqq 2$ and $\operatorname{Re}(A_z/B_z) \geqq -1$.

Then the inequality

$$\left| w^2(z) + A_z w(z) + B_z \frac{1+|w(0)|}{1-|w(0)|} z w'(z) \right| < 1,$$

for all $z \in D$, implies $|w(z)| < 1$, $z \in D$.

Coloring the Edges of a Graph

6034 [1975, 592]. *Proposed by Fred Galvin, University of California, Los Angeles*

Suppose the edges of the complete graph on n vertices are colored so that no color is used more than k times. (1) If $n \geqq k+2$, show that there is a triangle no two of whose edges are the same color. (2) Show that this is not necessarily so if $n = k+1$.

Solution by Allen J. Schwenk, U.S. Naval Academy. Let the edges of K_n be colored with no colors used more than k times.

(1) Let each color i induce a spanning subgraph G_i. Let $H \subset G_j$ denote a connected component with the most vertices appearing in any G_i. Subgraph H cannot have n vertices, for then there are at least $n-1 \geqq k+1$ lines of color j. Let u be any vertex not in H. If only one color, say m, joins u to H, then we have found a connected component of color m larger than H. Hence, at least two colors are used to join u to H. Since H is connected, this implies that there exist two points, v and w, adjacent to H joined by different colors to u. But that means u, v, w form a 3-colored triangle.

(2) For $n \leqq k+1$, let $v_1, v_2, \ldots, v_n$ denote the vertices of K_n. For $i < j$, we use color i to color the edge $v_i v_j$. Obviously, the color i is used exactly $n - i \leqq k$ times, and every triangle is 2-colored.

Also solved by B. Alspach & M. Gersom & G. Hahn & P. Hell (Canada), Andreas Blass, Aage Bondesen (Denmark), Michael Doob (Canada), Ulrich Faigle, O. P. Lossers (Netherlands), Ralph McKenzie, Walter Taylor (Australia), Joseph Weinstein, Dorothy Wolfe, and the proposer.

Note. Pavol Hell has generalized the problem and obtained the following: A coloring of the edges of the complete graph on n vertices is (k, l)-good if no color is used more than k times and no complete subgraph with l vertices has all $\binom{l}{2}$ edges of different color. Let $\operatorname{ar}(k, l)$ denote the smallest n such that the complete graph on n vertices does not admit a (k, l)-good coloring. Then the problem is the case $\operatorname{ar}(k, 3) = k+2$. The determination of other values of $\operatorname{ar}(k, l)$ appears to be difficult, but we note $k(l-1)+1 \leqq \operatorname{ar}(k, 1) \leqq \frac{1}{4}(k+1) \cdot l \cdot (l-1)(l-2)+3$. The lower bound is obtained by construction and the upper bound by counting the number of pairs of edges of the same color.

A Subseries of $\Sigma \mu(n) \log n/n$

6035 [1975, 529]. *Proposed by Arthur Marshall, Madison, Wisconsin*

For every natural number k let N_k be the kth number in natural order of the sequence consisting solely of primes and the (square-free) products of (two or more) successive primes. Let $\mu(\cdot)$ be the Moebius function: $\mu(N_k) = (-1)^r$, where r is the number of primes dividing N_k. Does the series

$$\sum_{k=1}^{\infty} \frac{\mu(N_k)}{N_k} \ln N_k$$

diverge (positively or negatively), converge, or oscillate?

[Note: It is known (Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, Vol. 2, pp. 585–587) that

$$\sum_{k=1}^{\infty} \frac{\mu(k)\ln k}{k} = -1.$$

Compare also Problem E 2258 [1971, 909]; also Segal, Proc. Amer. Math. Soc., 20 (1969), pp. 287, ff.; also Apostol, Proc. Amer. Math. Soc., 40 (1973), pp. 341, ff.]

Solution by Jeff Lagarias, Bell Laboratories, Murray Hill, New Jersey. Let $\mathcal{S} = \{2,3,5,6,7,\ldots\}$ denote the sequence of integers of the form $p_k p_{k+1} \cdots p_{k+r}$, that is products of any $r+1$ consecutive primes for $r = 0,1,2,\ldots$, arranged in increasing order, with the kth term denoted by n_k. Here p_k denotes the kth prime. Then the partial sums

$$S(N) = \sum_{n_k \leqq N} \frac{\mu(n_k)\log n_k}{n_k}$$

diverge to $-\infty$ as $N \to \infty$.

The key observation is that the contribution of terms with n_k a product of more than two primes is bounded. In fact it is obvious that $p_k p_{k+1} \cdots p_{k+r} > p_k^{r+1}$ while by use of Bertrand's postulate ([**1**] p. 110) that $p_{k+1} < 2p_k$ it follows that

$$p_k p_{k+1} \cdots p_{k+r} < 2^{r(r+1)} p_k^{r+1}.$$

Now we estimate, letting $\# n_k$ denote the number of distinct prime factors of n_k,

$$\sum_{\# n_k \geqq 2} \left| \frac{\mu(n_k)\log n_k}{n_k} \right| \leqq \sum_{\# n_k \geqq 2} \frac{(r+1)\log p_k + r(r+1)\log 2}{p_k^{r+1}}$$

$$\leqq \sum_p \left[p^{-2} \log p \sum_{r=0}^{\infty} \frac{(r+1)^2}{p^r} \right].$$

The inner sum is absolutely bounded independent of p, since

$$\sum_{r=0}^{\infty} \frac{(r+1)(r+2)}{p^r} \leqq \sum_{r=0}^{\infty} \frac{(r+1)(r+2)}{2^r} = C_0$$

whence

$$\sum_{\# n_k \geqq 2} \left| \frac{\mu(n_k)\log n_k}{n_k} \right| \leqq C_0 \sum_p p^{-2} \log p \leqq C_0 \sum_{n=1}^{\infty} n^{-2} \log n < \infty.$$

Hence the behavior of $S(N)$ depends solely on the behavior of

$$\bar{S}(N) = \sum_{p<N} \frac{\mu(p)\log p}{p} = -\sum_{p<N} \frac{\log p}{p}.$$

It is well known ([**1**] p. 101) that

$$-\bar{S}(N) = \sum_{p \leqq N} \frac{\log p}{p} = \log N + O\left(\frac{\log N}{\log\log N}\right)$$

diverges to $+\infty$, whence the result.

Reference

1. W. J. LeVeque, Topics in Number Theory, Vol. I, Addison-Wesley, 1956.

Also solved by L. E. Mattics, Gird Fricke, and Allen Stenger.

Even Perfect Numbers

6036 [1975, 671]. *Proposed by Carl Pomerance, University of Georgia*

If n is a natural number, we let $\sigma(n)$ denote the sum of the divisors of n, $S(n)$ the set of prime

divisors of n, and $\omega(n)$ the cardinality of $S(n)$. Clearly if n is an even perfect number, then $S(n)=S(\sigma(n))$ and $\omega(n)=2$. Prove the converse.

I. *Solution by David Melega, Freshman, University of Massachusetts.* If $\omega(n)=2$ then n has two distinct prime divisors, say A and B. Let $n=A^xB^y$, where $x \geqq 1$, $y \geqq 1$. If $S(n)=S(\sigma(n))$ then $\sigma(n)$ must equal $A^jB^kn=A^{j+x}B^{k+y}$, where $j \geqq 0$, $k \geqq 0$, and at least one of j and k is greater than zero (since $\sigma(n)\neq n$).

$$\frac{\sigma(n)}{n}=\frac{\sigma(A^x)\sigma(B^y)}{A^xB^y}=\frac{(A^{x+1}-1)(B^{y+1}-1)}{(A-1)(B-1)}\Big/(A^xB^y)$$

$$=\frac{(A^{x+1}-1)(B^{y+1}-1)}{A^xB^y}\cdot\frac{1}{(A-1)(B-1)}$$

$$=\frac{AB}{(A-1)(B-1)}-\frac{A^{x+1}+B^{y+1}-1}{n(A-1)(B-1)}=A^jB^k.$$

Now, since $A^{x+1}+B^{y+1}-1>0$, we have

$$\frac{AB}{(A-1)(B-1)}>A^jB^k \quad\text{and}\quad \frac{1}{(A-1)(B-1)}>A^{j-1}B^{k-1}.$$

Since the inequality is symmetric with respect to j and k, either $j=0$ or $k=0$. Let $k=0$, then

$$\frac{1}{(A-1)(B-1)}>\frac{A^{j-1}}{B} \quad\text{and}\quad \frac{1}{A^{j-1}(A-1)}>\frac{B-1}{B}.$$

Since $B \geqq 2$, $1/A^{j-1}(A-1)>\frac{1}{2}$, and $A^{j-1}(A-1)<2$. By trial and error we find that $j=1$ and $A=2$ are the only possible solutions. Therefore $n=2^xB^y$ and $\sigma(n)=2^{x+1}B^y=2n$, so that n is an even perfect number.

II. *Solution by Anton Glaser, Pennsylvania State University.* Let p and q be the two primes $(p<q)$ that are the divisors of both n and $\sigma(n)$. It follows that there exists an integer $k>1$ such that $k=p^aq^b$ and $\sigma(n)=kn$. This already establishes that n is pluperfect of multiplicity k. It is well known (Dickson, *History of the Theory of Numbers*, vol. I, pp. 33–38) that the multiplicity cannot exceed the number of prime divisors of n. This forces $k=2$, which means n is perfect. Furthermore, $k=2$ implies $p=2$ $(a=1;\ b=0)$ and hence, n is also even.

Also solved by David Bressoud, Robert Breusch (New Zealand), Hugh Edgar, Robert Gilmer, M. G. Greening (Australia), O. P. Lossers (Netherlands), Arthur Marshall, L. E. Mattics, T. Šalát (Czechoslovakia), K. R. Pratap Singh (India), E. W. Trost (Switzerland), University of British Columbia, Grads, Problem Group, and the proposer.

Imbedding of a Graph

6037 [1975, 671]. *Proposed by Jim Lawrence, University of Washington*

Show that any graph H is isomorphic to an induced subgraph of some finite graph H' which has a group of automorphisms that acts transitively on its vertices.

Solution by Allen J. Schwenk, U.S. Naval Academy. The given graph H has vertices u_0, $u_1,\ldots,u_{p-1}$. We shall construct a graph H' whose automorphism group is transitive on its point set and which has H as an induced subgraph. We arrange $n=2^p-2$ points in a circle and label them modulo n by $v_1, v_2, v_3,\ldots, v_n=v_0$. For each edge u_iu_j in H with $i<j$, we insert in H' for each k the edge v_kv_{k+r} where $r=2^j-2^i$. As constructed, the dihedral group of order $2n$ must be a subgroup of the automorphism group $\Gamma(H')$, and so $\Gamma(H')$ is transitive on the vertices. Furthermore, for each value of r, the subgraph induced by $v_{1+r}, v_{2+r}, v_{4+r},\ldots, v_{2^{p-1}+r}$ is isomorphic to H.

Also solved by the proposer.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN

with the assistance of the mathematics departments of St. Olaf and Carleton Colleges

COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, CARLETON COLLEGE

We invite readers to submit reviews of significant recent college-level mathematics books. We especially encourage reviews based on classroom use, or comparative reviews of several related books. Reviews should ordinarily not exceed two pages (per book) typed double spaced. Manuscripts of reviews as well as books submitted for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057.

Mathematics Foundation Course. By the Open University Mathematics Foundation Course Team. The Open University Press, Walton Hall, Bletchley, Bucks, England, 1971. Distributed by Open University Educational Enterprises, Inc., North American Office, Washington, D.C.

The Open University of Great Britain provides an extensive external degree program for adult students. A typical course, which is equivalent to about 18 semester credits in the United States, is taken by thousands of students living throughout Great Britain. Each week the students study a 40–60 page booklet, listen to a 30-minute radio lecture, watch a 25-minute BBC television lecture, mail in an exercise set to be graded, and can meet with a tutor at one of the approximately 260 study centers located throughout Great Britain. In addition, students attend an intensive one-week residence course, usually in the middle of the 10-month term.

For the past three years the University College division of the University of Maryland has been using the materials developed at the Open University for their beginning course in mathematics called the Mathematics Foundation Course. These materials consist of 32 booklets and exercise sets, which the student purchases, and 36 sets of audio-tapes and films which are used in class. Students work primarily at home and can attend one three-hour session each week for tutorials and the viewing of the films.

The Mathematics Foundation Course is an integrated course dealing with the calculus, linear algebra, probability and statistics, numerical analysis, computing (omitted at the University of Maryland), logic, complex numbers, modern algebra, and topology. In each area an attempt is made to present the basic ideas and to stress those concepts which cause the greatest difficulty in upper division courses (e.g., morphisms, limits, linear independence). The units of the course are arranged in a spiral approach to give students two or three weeks in which to absorb a new concept before using the concept again. For example, the sequence leading up through the second unit of integral calculus is the following: functions, errors and accuracy, morphisms and operators, finite differences, inequalities, sequences and limits, computing I, integration I, logic–boolean algebra, differentiation I, integration II.

As it is currently designed the course gives the student a look at a large portion of the material an undergraduate mathematics major would study. But in the process of achieving this breadth of outlook considerable depth is lost. For example, only a very few applications of the integral or the derivative and a few methods of solving differential equations are included. In the units on probability and statistics there is little discussion of the role of variance or standard deviation (a fact which disturbed many of our students) and no work with continuous distributions although students do have the calculus available by that time. Also excluded in general are proofs of theorems.

Students who complete the course should find it a good preparation for upper division mathematics courses, particularly courses such as abstract algebra in which the calculus does not play a major role. In addition there are Open University advanced courses which are designed specifically to follow this foundation course.

A major problem for our own adult students has been the lack of time to brush up on their often

rusty algebra or trigonometry skills or to develop any ability to handle more theoretical concepts. Although review booklets are available, the reading of 40–60 pages every week leaves little opportunity for review of such essential topics as trigonometry, graphing or solutions of systems of linear equations. Since many students have not worked with mathematics for several years, such a review is necessary. (We are now in the process of revising the University of Maryland course to allow two weeks at the beginning of the first term for study of essential pre-calculus topics.)

Even without the rustiness in mathematical skills many students find the course unusually demanding in its first weeks simply because of the theoretical nature of the units on errors and accuracy and on morphisms and operators (particularly the latter since it deals with induced binary operations to the extent that students forget that they are looking for "morphisms" — homomorphisms and isomorphisms). In our revision of the course we have placed these two topics—in a rewritten form—together with the two units on integration after the first unit on differentiation.

Considerable effort has been taken to distinguish between a function and the value of the function. This has resulted in an atypical notation not only for the function but for the derivative and integral of a function. For instance, the squaring function is written $x \to x^2$ $(x \in \mathbf{R})$ while its derivative is denoted $D(x \to x^2) = x \to 2x$ $(x \in \mathbf{R})$ and the integral as $\int x \to x^2 = x \to (1/3)x^3$ $(x \in \mathbf{R})$. While this notation may help the student initially we feel that it has the drawback of making certain topics (e.g., substitution in integration) unnecessarily cumbersome. In addition it is difficult if not impossible for most students to use other calculus texts as references because of the difference in notation. While this will not stop us from using the foundation course at the University of Maryland we do hope that the revision which is currently underway at the Open University will use a more standard notation, at least for the integral and derivative.

Many of our students have found the course interesting and challenging but most of those who have done well in the course have also had as much as a complete course in the calculus before starting the foundation course. Those students who have concerned us the most are the ones who might have completed the course if we could have given them time for pre-calculus topics and eased the beginning of the course by moving the theoretical units to a later part of the course. In fact, our experience with the course in the Fall 1975 term has been that students have remained interested and have been able to stay on schedule with assignments now that we have rearranged the topics and added two weeks of pre-calculus work. We look forward to the revision of the course which is being written now by the Open University and expect that the experience of the Open University with their own classes will lead to changes in many of the areas where we have had problems.

There is no question that the Open University Mathematics Foundation Course is a contribution to rethinking the goals and content of the introductory program for mathematics majors, particularly for the adult student. Its considerable success in Great Britain, where over ten thousand students have completed the course and large numbers of students have gone on successfully to study upper level mathematics in the Open University, attests to its strengths. Although we doubt that this course will ever be presented to large numbers of students in the United States, in part because of the mathematical maturity and the amount of work required, we feel that its innovative content and carefully thought-out presentation generally outweigh its shortcomings. We certainly plan to continue offering the course in the future.

FRANCES GULICK AND STEPHEN WRIGHT, University of Maryland, University College

Mathematics for the Biological Sciences. By Stanley I. Grossman and James E. Turner. Macmillan, New York, 1974. xi + 512 pp. $11.50. (Telegraphic Review, October 1974.)

Having used this text in a course for biology students, I believe that the authors have fulfilled their objectives: to make relevant mathematics accessible in a reasonable amount of time; and to develop the student's ability to relate mathematics to problems in biology and medicine.

Written clearly and carefully with a wealth of examples and well-posed exercises, the book requires a certain maturity and sophistication of its readers. There is ample material in this text for a three-hour, one semester course in non-calculus mathematics. The non-calculus topics include: discrete probability, Markov chains, vectors and matrices, linear programming, game theory and difference equations. There is also a chapter on biological models which primarily uses material on difference equations.

Some features of this textbook which I found especially attractive were its speedy treatment of sets and functions, absence of symbolic logic, inclusion of the Poisson distribution and of systems of first order difference equations, and the treatment of the linear programming problem $Ax \leqq b$ where b has negative entries. In addition to these the solution of matrix games by linear programming methods and the selection and fitness models were treated.

The text also includes chapters on differential equations and on continuous probability. This of course requires some calculus. The former includes material on systems of first order differential equations which is used in deriving a mathematical model for competing species. The latter includes the usual discussions on the normal distribution and confidence intervals.

If the instructor prefers he can, without loss of continuity, easily delete the material on linear programming and game theory, in favor of some of the calculus-based mathematical topics.

In summary, I recommend that anyone faced with teaching a mathematics course to biological science students consider this text.

JOHN J. BUONI, Youngstown State University

Statistics. By Norma Gilbert. W. B. Saunders Co., Philadelphia, Pennsylvania, 1976. ix + 364 pp. $12.95. (Telegraphiç Review, March 1977.)

This text is a well-written, freshman-sophomore level treatment aimed at students who have little mathematical background. The first chapter is a review of most of the basic mathematics needed in a non-calculus course in statistics. Other chapters cover all of the traditional topics from an intuitive viewpoint, including probability and nonparametric tests.

The author attempts to give the reader an understanding of the many difficult concepts in so-called "elementary statistics" through the use of a concise, convincing discussion interspersed with an abundance of examples. In my opinion she succeeds admirably. The presentation does not bog down with precise definitions when intuitive ones will do, and at the same time, it does not convey false ideas.

Heavy black print is used to point out important concepts and color is used for important formulas. This makes the text particularly readable. At the end of each chapter there is a section which lists the important words and symbols contained in the chapter. The text has a wealth of exercises with answers to as well as explanations of all odd numbered exercises given at the end of each chapter.

The author has made a few bad choices of symbols, e.g. the same symbol is used for the probability of an event, for the probability of a success in a Bernoulli trial and for a sample proportion. There are also a fair share of first printing errors which it is to be hoped will be corrected in a second printing.

The reviewer has already used the text in a summer course and has adopted it for the coming school year. Overall, I believe it to be one of the very best of its kind to appear in recent years.

LELAND D. GRABER, Central College

TELEGRAPHIC REVIEWS

Telegraphic reviews are designed to give prompt notice of new books with sufficient information to assist our readers in deciding whether to order an examination copy or to suggest library purchase. Possible uses are indicated as follows:

T = textbook P = professional reading
S = supplementary reading L = undergraduate library purchase
13 to 18 = freshman to second year graduate level usage
1 to 4 = appropriate time in semesters to cover text

Asterisks (*) or question marks (?) denote special positive or negative emphasis, respectively. Publishers are denoted by standard abbreviations; complete addresses may be found in *Books in Print.*

GENERAL, P*, L. *Undergraduate Mathematical Sciences in Universities, Four-Year Colleges and Two-Year Colleges, 1975-76.* Ed: James T. Fey, Donald J. Albert, John Jewett. CBMS, 1976, xii + 130 pp, $4 (P). The third published report in a series of four successive nationwide statistical surveys of collegiate education conducted at five-year intervals (1960, 1965, 1970, 1975). Covers student, faculty and curricular data from four-year and two-year institutions, with comparison (where appropriate) to previous surveys. Confirms and quantifies many well known patterns (e.g., the precipitous drop in enrollments in upper division mathematics courses) and reveals some surprises (e.g., slide rule courses still outdraw hand calculator courses in two-year institutions). An important and unique reference for chairmen and others concerned with the future of collegiate mathematics education. LAS

GENERAL, S. *Standard Mathematical Tables, Twenty-fourth Edition.* Ed: William H. Beyer. CRC Pr, 1976, 565 pp, $9.95. 20% less material than in recent editions; omissions make the volume lighter, not significantly less useful. LAS

BASIC, T(13: 1). *Beginning Technical Mathematics.* Jerry O'Donnell. Reston, 1976, xi + 250 pp, $12.95. Elementary algebra developed around technical problems. Addition of fractions to quadratic equation. Reading sine from the slide rule. Decibel and ph-scales as examples of use of logarithms. LH

BASIC, T(13: 1). *Arithmetic, A Word, Diagram, Symbol Approach.* Merlin Miller, John Walsh. Macmillan, 1976, x + 422 pp, $10.95 (P). Operations with whole numbers, fractions, decimals, signed numbers. Stress on understanding problems from three points of view: verbal, symbolic and visual (e.g., via the number line). Exercises on perforated worksheets follow examples. Section and chapter review exercise sets. LCL

PRECALCULUS, T(13: 1). *Essentials of Algebra for College Students.* Raymond A. Barnett. McGraw, 1977, xvi + 382 pp, $11.95. Clear exposition with sufficient examples. Numerous exercises grouped into three levels of difficulty. Answers to odd problems. Has available supplements for use in self-paced instruction. Discussion of polynomials is limited (e.g., no synthetic division). One very nice feature is application sections for many of the chapters with numerous exercises. Covers standard precalculus algebra including exponentials and logs. CB

PRECALCULUS, T*(13: 1). *Essentials of Precalculus Mathematics.* Dennis T. Christy. Har-Row, 1977, viii + 568 pp, $13.95. Kudos to Dennis Christy for using some ingenuity in writing a text which is not "just another precalculus text." There is an early introduction to trigonometry which makes it possible to include trigonometric functions in factoring, etc. Returns to trigonometric functions of real numbers later. Also introduces such topics as use of logarithmic paper, discussion of musical sounds and radio waves. LLK

EDUCATION, T(15-16: 1). *Geometry for Teachers.* C. Patrick Collier. HM, 1976, xiv + 331 pp, $12.50. For elementary teacher education. Considers informal geometry and topology, motion geometry and metric geometry. Excellent pedagogy apparent in a new instructional plan involving information giving, search activities and quizzes. Worth considering. PSJ

EDUCATION, S(15-16), P. *Activities Handbook for Teaching the Metric System.* Gary G. Bitter, Jerald L. Mikesell, Kathryn Maurdeff. Allyn, 1976, vi + 378 pp, $12.95. A resource for workshops, inservice courses, individual study, or the elementary and secondary classrooms. Provides more than 100 activities involving exploration, discovery and drill and practice. Topics include linear measure, volume and capacity, area and perimeter, temperature and mass. Appendices offer diagnostic and mastery tests, bibliography, tables and metric supplies. PSJ

EDUCATION, T(15-16: 1), P. *Elementary School Mathematics: Teaching the Basic Skills.* William Zlot. T.Y. Crowell, 1976, xiv + 494 pp, $13.50. A methods text. Considers cardinal numbers, fractional numbers, and the four fundamental operations. Discusses many teaching methods and instructional devices that a teacher may use to supplement the classroom text. Topic orientation rather than grade level. Excellent bibliographies with each chapter supported by lists of bibliographical activities. Instructional objectives and exercise sets at end of each chapter. PSJ

EDUCATION, S(15), P, L. *Teaching Metrics Simplified.* James B. Cunningham. P-H, 1976, viii + 184 pp, $10.95; $5.95 (P). For elementary and secondary teachers. Comprehensive treatment of SI for those with little previous knowledge of metric measure. Activity oriented approach stresses "thinking metric" via extensive sets of general and supplementary classroom activities. PSJ

EDUCATION, P. *Beiträge zum Mathematikunterricht 1976.* Hermann Schroedel, 1976, 286 pp, DM 21,80 (P). Proceedings of the tenth (1976) conference on mathematics education. Contains a large number and wide variety of papers of interest at least through the junior-college level. JAS

HISTORY, P, L*. *Graph Theory, 1736-1936.* Norman L. Biggs, E. Keith Lloyd, Robin J. Wilson. Clarendon Pr, 1976, xi + 239 pp, $21. An engaging historical survey. Concentrates on 10 topics including paths, circuits, trees, chemical graphs, Euler's polyhedral formula, coloring problems, and

factorization of graphs. The exposition revolves around extracts from 37 articles from Euler to Hamilton to Whitney. Experts and nonexperts will find this book both informative and entertaining. SG

HISTORY, P, L. *The New Elements of Mathematics, V. III-IV.* Charles S. Peirce. Ed: Carolyn Eisele. Humanities Pr, 1976. *V. III/1: Mathematical Miscellanea,* xxxix + 763 pp; *V. III/2,* 386 pp, $128; *V. IV: Mathematical Philosophy,* xxviii + 393 pp, $48. The final volumes of Peirce's mathematical papers. (The first two volumes were published earlier; see TR, November 1976.) Each volume contains an introduction by the editor. LAS

FOUNDATIONS, T(17-18), P. *Set Theory with an Introduction to Descriptive Set Theory.* K. Kuratowski, A. Mostowski. Stud. in Logic and Found. of Math., V. 86. North-Holland, 1976, xiv + 514 pp, $29. Major changes from the 1967 edition include a new chapter on trees (with a short introduction to the partition calculus), a complete rewrite of the chapter on inaccessible cardinals, and 4 new chapters (by Kuratowski) on descriptive set theory (including original results) to replace the former Chapter 10 on analytic and projective sets. The framework of classical set theory is maintained--no arguments of a model-theoretic nature appear. PJC

COMBINATORICS, T(15-17: 1). *An Introduction to Algebraic and Combinatorial Coding Theory.* Ian F. Blake, Ronald C. Mullin. Acad Pr, 1976, xi + 229 pp, $9.50 (P). A paperbound edition of chapters 1-3 of *The Mathematical Theory of Coding* by the same authors. The first chapter covers finite fields, linear codes, the second discusses codes arising from finite geometries, designs, difference sets, quadratic residues. Matroids and perfect codes are among the topics in chapter 3. For an extended review of the original text see *Bull. Amer. Math. Soc.,* 82 (1976) 41-42. SG

COMBINATORICS, T(13-15: 1), L. *Dots and Lines.* Richard J. Trudeau. Kent St U Pr, 1976, x + 199 pp, $11; $6.50 (P). An introduction to graph theory intended for both non-math and math students. Topics: graphs, planar graphs, platonic graphs, Euler's theorem, coloring, genus, Eulerian and Hamiltonian graphs. SG

NUMBER THEORY, P. *Trigonometrical Sums in Number Theory.* I.M. Vinogradov. Statistical Pub, 1975, viii + 152 pp, $15. An exposition of important results and applications of trigonometrical sums: Weyl's method, Waring's problem, sums over prime numbers. Most of the discussion concerns the author's own work. SG

LINEAR ALGEBRA, T*(14-15: 2), L. *Linear Algebra.* Nathan Divinsky. Page Ficklin, 1975, xxxx + 314 pp, $12.95. Very readable exposition. Good examples. Uses identification of quadric surfaces as motivation for developing algebraic tools. Matrix theory nicely intermixed with linear algebra. Number of problems is on the sparse side. Only one mention of outside applications. CB

LINEAR ALGEBRA, T(14-15: 1). *Linear Algebra.* Norman J. Bloch, John G. Michaels. McGraw, 1977, x + 342 pp, $12.95. A rather gentle introduction, assuming little previous knowledge or sophistication. Chapters on linear equations, vector spaces (abstract but real) and subspaces, bases and dimension, linear transformations, characteristic values and vectors, and Euclidean geometry in vector spaces. Applications and review questions in every chapter. JD-B

ALGEBRA, T(18), P*. *Algebra II: Ring Theory.* Carl Faith. Gund. math. Wissenschaften, B. 191. Springer-Verlag, 1976, xvi + 302 pp, $40.20. A major resource work covering aspects of ring theory since the publication of Jacobson's *Structure of Rings.* As in the first volume, much literature is quoted by means of referenced exercises. A third volume *Commutative Rings, Hereditary Rings, Separable Algebras and the Brauer Group,* is forthcoming and will complete the series. Note price! LCL

ALGEBRA, T(15: 2). *Algebra.* L.E. Sigler. Springer-Verlag, 1976, xi + 419 pp, $14.80. Leisurely introduction: sets, rings (four chapters--basics, integers, rationals, polynomials), linear algebra (three chapters), monoids and groups (only one short chapter!). Principles of universal algebra are incorporated--particularly with regard to consistent terminology and notation. Multiple choice questions among the exercises provide a quick check on understanding of definitions. LCL

ALGEBRA, T(15-16: 1), S. *Foundations of the Theory of Groupoids and Groups.* O. Borůvka. Trans: M. Borůvkova. Halsted Pr, 1976, 215 pp, $24.75. Translation of 1960 German edition. Has three parts, of equal length: sets, groupoids, groups. Content of the first and third is not uncommon to undergraduate modern algebra courses. Groupoids are studied using decompositions into sets and using homorphic mappings. Exercise sets do not seem comprehensive or very challenging. DFA

ALGEBRA, T(17-18), S, P*, L. *Rings with Involution.* I.N. Herstein. U of Chicago Pr, 1976, x + 247 pp, $5.50 (P). A pleasurable sampler typlifying the work currently being done in the area of rings with an involution; stops short of applications to the theory of Jordan algebras, operator algebras, and Banach algebras. Self-contained modulo the author's *Topics in Ring Theory* (TR, March 1972). LCL

CALCULUS, S(13). *Problemas y Ejercicios de Análisis Matemático.* Ed: B. Demidovich. MIR, 1977, 519 pp. Compendium of more than 3000 calculus exercises (with answers and some complete solutions), organized by topic. Each section begins with a brief theoretical introduction, with definitions, important formulas, and examples. Translation into Spanish of the fifth Russian edition. PJC

COMPLEX ANALYSIS, P. *Lecture Notes in Mathematics-538: Complex Analytic Geometry.* Gerd Fischer. Springer-Verlag, 1976, vii + 201 pp, $9.50 (P). "A survey of the fundamental concepts and results of...the theory of functions of several complex variables... ." Large bibliography. CB

DIFFERENTIAL EQUATIONS, P. *Lecture Notes in Mathematics-535: Singularités d'Applications Différentiables.* Ed: O. Burlet, F. Ronga. Springer-Verlag, 1976, 253 pp, $10.30 (P). Proceedings of the seminar held at Plans-sur-Bex in 1975. JAS

DIFFERENTIAL EQUATIONS, T(14-15). *Elementary Differential Equations and Boundary Value Problems, Second Edition.* William E. Boyce, Richard C. DiPrima. Wiley, 1969, xiv + 584 pp, $16.95. A new printing of the second edition, containing a new chapter on stability theory for linear and nonlinear equations and a reorganization of the material on systems of equations and partial differential equations, Fourier series and eigenvalue problems. This thorough introductory text contains many applications, examples and tons of exercises. SG

DIFFERENTIAL EQUATIONS, T(18: 1, 2), P. *Direkte Methoden der Variationsrechnung.* Waldemar Veite. Teubner, Stuttgart, 1976, 198 pp, DM 24,80 (P). A straightforward exposition of the theory of variational methods for partial differential equations, including some background in linear spaces and operators. JAS

DIFFERENTIAL EQUATIONS, P. *Asymptotic Methods and Singular Perturbations.* Ed: Robert E. O'Malley, Jr. SIAM-AMS Proc., V. X. AMS, 1976, vi + 154 pp, $16.40. Proceedings of the symposium in applied mathematics of the AMS and SIAM held in New York City in April 1976. JAS

DIFFERENTIAL EQUATIONS, S(16-17). *Les Équations aux Dérivées Partielles, en Physique et en Mécanique des Milieux Continus.* Serge Colombo. Masson, 1976, viii + 188 pp, 90F. Aimed at physicists and engineers. Emphasis on equations arising from physical problems: diffusion equations, Laplace equation, hyperbolic equations. Also contains exposition of basic results on partial differential equations. Some exercises included. SG

NUMERICAL ANALYSIS, P. *Moderne Methoden der Numerischen Mathematik.* J. Albrecht, L. Collatz. Int. Ser. Num. Math., V. 32. Birkhauser, 1976, 175 pp, sFr. 34 (P). Proceedings of the conference held in June 1975 at Clausthal in honor of the two-hundredth anniversary of the University of Clausthal. JAS

NUMERICAL ANALYSIS, P. *Numerische Behandlung von Differentialgleichungen, Band 2.* J. Albrecht, L. Collatz. Int. Ser. Num. Math., V. 31. Birkhauser, 1976, 276 pp, sFr. 44 (P). Proceedings of a conference emphasizing the method of finite elements which was held in Oberwolfach in November 1975. JAS

NUMERICAL ANALYSIS, T(18: 1), S, P. *An Introduction to the Mathematical Theory of Finite Elements.* J.T. Oden, J.N. Reddy. Wiley, 1976, xii + 429 pp, $24.95. The first half of the book is on mathematical preliminaries: distributions, mollifiers, generalized derivatives, Hilbert and Sobolev spaces; the second half is the theory behind applications to interpolation, variational boundary-value problems, elliptic problems and time-dependent problems. Prerequisite: some functional analysis. Bibliographies. RWN

NUMERICAL ANALYSIS, P, L. *Modern Numerical Methods for Ordinary Differential Equations.* Ed: G. Hall, J.M. Watt. Clarendon Pr, 1976, 336 pp, $21.50. A broad survey of and guide to the theory and practice of computational methods for the solution of differential equations. Includes initial-value, stiff, boundary value, eigenvalue, delay-differential and integro-differential problems. A 22-page bibliography is referenced liberally throughout the text. Authored primarily by researchers at the Universities of Manchester and Liverpool. A good book to which to refer first. RWN

FUNCTIONAL ANALYSIS, P. *Functional Analysis.* Ed: Djairo Guedes de Figueiredo. Lect. Notes in Pure and Appl. Math., V. 18. Dekker, 1976, viii + 325 pp, $24.50 (P). Proceedings of a July 1974 symposium held at Campinas, Brazil; contains nineteen invited and contributed papers. LAS

FUNCTIONAL ANALYSIS, S(18), P. *Nonlinear Analysis.* Ed: N.M. Temme. Math Centrum, 1976. *V. 1*, viii + 135 pp, Dfl. 15 (P); *V. 2*, viii + 198 pp, Dfl. 21 (P). Proceedings of the colloquium lectures on nonlinear analysis held at the Mathematisch Centrum (Amsterdam) in 1975. Theory and applications of topological degree and nonlinear eigenvalue problems are presented in volume 1 for finite dimensional spaces and in volume 2 for infinite dimensional spaces. According to the editor, all results in these two volumes can be found in the literature. I-CH

FUNCTIONAL ANALYSIS, P. *Selected Problems of Weighted Approximation and Spectral Analysis.* N.K. Nikol'skii. Proc. of Steklov Inst. of Math., No. 120. AMS, 1976, iii + 276 pp, $37.60 (P). Concerned with the general problem of spectral "analysis-synthesis", this monograph contains selected problems which originate in classical harmonic analysis, weighted approximation by polynomials, and invariant subspaces of linear operators. Extensive bibliography. No index. Translated from Russian by F.A. Cezus. I-CH

FUNCTIONAL ANALYSIS, P. *Nonlinear Functional Analysis and Differential Equations.* Ed: Lamberto Cesari, Rangachary Kannan, Jerry D. Schuur. Lect. Notes in Pure and Appl. Math., V. 19. Dekker, 1976, xii + 352 pp, $29.50 (P). Proceedings of a conference held at Michigan State University in June 1975. The first half consists of the principal lectures delivered by L. Cesari. The second half collects the talks given by other participants, on topics ranging from problems in bifurcation theory to measurability of solutions of differential equations to current aspects of degree theory. I-CH

FUNCTIONAL ANALYSIS, S(18), P. *Ordered Vector Spaces and Linear Operators.* Romulus Cristescu. Editura Academiei Romania, 1976, 339 pp, Lei 30. Survey which contains recent work but begins with review of basic facts concerning lattices, topological vector spaces. Includes operators on topological vector lattices, integrals with respect to vector measures. LH

FUNCTIONAL ANALYSIS, P. *Produse Tensoriale Topologice si Bornologice.* Nicolae Popa. Editura Academiei Romania, 1976, 258 pp, Lei 12,50 (P).

OPTIMIZATION, P. *Stochastic Linear Programming.* Peter Kall. Springer-Verlag, 1976, 95 pp, $15.60. The theory and the methods of linear programming under uncertainty. The data considered (e.g., demands, technological coefficients, available capacities, cost rates, etc.) are random variables. The author successfully chooses and presents those topics and results which can be handled more or less systematically within a certain theoretical framework. I-CH

OPTIMIZATION, T*(17-18: 2), S*. *Nonlinear Programming, Analysis and Methods.* Mordecai Avriel. P-H, 1976, xv + 512 pp, $24.95. The presentation includes many ideas from the author's own research papers as well as many topics usually not found in other volumes of this type--conjugate function theory, duality in nonlinear programming, unified approach to variable metric algorithms, etc. It is modern nonlinear programming, much of it at the frontiers, done in modern ways. I-CH

OPTIMIZATION, S(15-17). *Classical Optimization: Foundations and Extensions.* Michael J. Panik. Stud. in Math. and Managerial Econ., V. 16. North-Holland, 1976, xi + 312 pp, $16.95 (P). Following many pages devoted to "foundations" (most of them also belonging to standard single or multi-variable calculus or matrix theory), the book addresses its second half to some well-known methods of constrained optimization. Treats techniques of Lagrange multipliers and Kuhn-Tucker theory in detail. Many illuminating geometric interpretations and illustrations. No exercises. I-CH

OPTIMIZATION, P. *Optimization in Action.* Ed: L.C.W. Dixon. Acad Pr, 1976, xi + 622 pp, $29.50. Proceedings of the conference held at the University of Bristol, in January 1975. JAS

OPTIMIZATION, S(18), P. *Lecture Notes in Mathematics-479: Minimum Norm Extremals in Function Spaces, with Applications to Classical and Modern Analysis.* Stephen D. Fisher, Joseph W. Jerome. Springer-Verlag, 1975, viii + 209 pp, $9.90 (P). Functional analysis applied to existence and characterization of solutions in approximation theory, differential equations and control theory. Optimal controls for systems governed by differential equations and inequality constraints, minimum curvature problem, perfect spline functions as extremals. LH

OPTIMIZATION, S(17-18), P. *Lecture Notes in Economics and Mathematical Systems-110: Optimal Control of Discrete Time Stochastic Systems.* Charlotte Striebel. Springer-Verlag, 1975, 208 pp, $9.90 (P). Self-contained treatment, emphasizes sufficient conditions for optimality. Aims to develop algorithms for the construction of optimal control laws. LH

OPTIMIZATION, T*(17-18: 1, 2), S*. *Nonlinear Programming for Operations Research.* Donald M. Simmons. P-H, 1975, xiii + 448 pp, $18.95. Contains light yet sufficient coverage of the theory. Purposely omits the algorithmic convergence theory to place the emphasis on practical methods. Presents those solution algorithms that are popular, and, in most cases, capable of fast convergence, simple to work with and to code for computer use. A good rigorous book, written in terse language by an experienced practitioner of operations research. I-CH

OPTIMIZATION, S(18), P. *Singular Optimal Control Problems.* David J. Bell, David H. Jacobson. Math. in Sci. and Eng., V. 117. Acad Pr, 1975, xi + 190 pp, $15. A presentation of some new theorems developed in the last decade for certain special control problems (called singular problems) in which the Pontryagin principle yields no additional information on the stationary control: new necessary conditions for singular extremals to be candidate optimal arcs, and new necessary and sufficient conditions for such extremals to be optimal. I-CH

ANALYSIS, T(16-18), S**, P**, L**. *Introduction to the Theory of Infinitesimals.* K.D. Stroyan, W.A.J. Luxemburg. Pure and Appl. Math., V. 72. Acad Pr, 1976, xiii + 326 pp, $24.50. A really thorough and charming exposition of Robinson's infinitesimals. The first part of the development covers much more than just a foundation for analysis--logic, rings, and even categories. Standard topics in non-standard calculus end the first part. The second part covers much more advanced topics in functional analysis. The inclusion of history and heuristics in a really readable way make this a tempting choice for a text if you have imaginative students and the time to make up appropriate problems. JAS

ANALYSIS, T(17: 2). *Analysis, Part I: Elements.* Krzysztof Maurin. Trans: Eugene Lepa. Reidel, 1976, xiii + 430 pp, $39. First of three volumes intending to lead from beginnings to frontiers; the second volume is *Integration, Distributions, Holomorphic Functions,* the third, *Introduction to Global Analysis. Elements* gets from single-variable calculus to ordinary differential equations and curves in a Banach space. DFA

ANALYSIS, T(18), P. *Operational Methods.* V.P. Maslov. Trans: V. Golo, N. Kulman, G. Voropaeva. MIR, 1976, 559 pp. Chapter titles: Introduction to Operational Calculus, Functions of a Regular Operator, Calculus of Noncommutative Operators, Asymptotic Methods, Generalized Hamilton-Jacobi Equations, Canonical Operator on a Lagrangean Manifold with a Complex Germ and Proof of the Main Theorem. Well-known methods and a new one, that of ordered operators. Theory, concrete applications, problems. DFA

ANALYSIS, P. *Lecture Notes in Mathematics-543: Nonlinear Operators and the Calculus of Variations.* Ed: J.P. Gossez, *et al.* Springer-Verlag, 1976, 237 pp, $10.30 (P). Lecture notes for the five series of lectures at the summer school held at the Université Libre de Bruxelles, in September 1975. Seminar results are not included. JAS

ANALYSIS, P. *Scattering Theory for Automorphic Functions.* Peter D. Lax, Ralph S. Phillips. Annals of Math. Stud., No. 87. Princeton U Pr, 1976, x + 300 pp, $18.50; $7.50 (P). A reformulation in terms of scattering theory of harmonic analysis of Sl(2,R). The development yields the spectral theory of the Laplace-Beltrami operator to the Selberg trace formula among other things. Should be of great interest to both analysts and number theorists. SG

ANALYSIS, T(15-16: 2). *Applied Mathematics Series: Mathematics for Engineering and Science.* H. Guggenheimer. Krieger, 1976, ii + 285 pp, $9.95 (P). For "applied advanced calculus" courses. Chapters on vectors, linear algebra, vector calculus, ordinary differential equations, Fourier series and integrals, partial differential equations. Uses linear algebra throughout. Includes numerical methods for hand calculators. Appears to need some fleshing out. Printed directly from typescript. DFA

ANALYSIS, P. *Lecture Notes in Mathematics-541: Measure Theory.* Ed: A. Bellow, D. Kölzow. Springer-Verlag, 1976, xiv + 430 pp, $14.40 (P). Proceedings of the conference held at Oberwolfach in June 1975. (Editor Alexandra Bellow was formerly A. Ionescu Tulcea.) JAS

ANALYSIS, P. *Lecture Notes in Mathematics-523: Mathematical Theory of Feynman Path Integrals.* Sergio A. Albeverio, Raphael J. Høegh-Krohn. Springer-Verlag, 1976, v + 139 pp, $7.50 (P). Aim is to "develop from scratch a self-contained theory of oscillatory integrals [on real Hilbert spaces]" and "apply it to the mathematical foundation of the...Feynman path integrals" of quantum mechanics. Extensive references. CB

ANALYSIS, P. *Lecture Notes in Mathematics-544: On the Functional Equations Satisfied by Eisenstein Series.* Robert P. Langlands. Springer-Verlag, 1976, vii + 337 pp, $12.80 (P). Development of functional equations for Eisenstein series together with appendices covering background material on Dirichlet series, adèle groups and examples. SG

ALGEBRAIC GEOMETRY, P. *Algebraic Geometry I, Complex Projective Varieties.* David Mumford. Grund. math. Wissenschaften, B. 221. Springer-Verlag, 1976, x + 186 pp, $14.80. An introduction to classical algebraic geometry. Topics include affine and projective varieties, correspondences, Chow's theorem, degree, linear systems, genus of curves, and the 27 lines on a cubic surface. Volume II will deal with schemes. SG

GEOMETRY, P. *Unitary Representations of Maximal Parabolic Subgroups of the Classical Groups.* Joseph A. Wolf. Memoirs No. 180. AMS, 1976, iii + 193 pp, $8.40 (P). A determination of the irreducible unitary representations of the maximal parabolic subgroups of the real and complex classical Lie groups.. SG

GEOMETRY, P. *Homogeneous Manifolds with Negative Curvature, Part II.* Robert Azencott, Edward N. Wilson. Memoirs No. 178. AMS, 1976, iii + 102 pp, $7.20 (P). The second in a series dealing with the structure of the full isometry group of a connected, simply connected, homogeneous, Riemannian manifold with non-positive sectional curvature. JAS

GEOMETRY, P. *Poristic Theory, Pedal Lines, Kantor Lines, Kantor Points and Allied Topics.* Kesiraju Satyanarayana. Visalaandhra Pub House, 1976, xviii + 157 pp, Rs. 30. A classical study of the topics listed in the title using analytic geometry with complex numbers, this volume is a limited printing (200 copies) of the 1964 "stencilled" version with a similar title. JAS

TOPOLOGY, P. *Analiza pe Corpuri Ultrametrice.* Gheorghe Isac, Gheorghe Marinescu. Editura Academiei Romania, 1976, 192 pp, Lei 19,50. A thorough presentation of the theory of ultrametric fields as used in point set topology and functional analysis. JAS

STATISTICS, T(13: 1). *Statistics.* Norma Gilbert. Saunders, 1976, ix + 364 pp, $12.95; *Study Guide for Statistics,* vii + 166 pp, $3.95 (P). Standard coverage (including probability, hypothesis testing, analysis of variance, correlation, nonparametrics) for students with minimal math background (high school algebra). Minimal notation; stress on understanding formulas. Study guide (complete with audio-tapes if desired) gives 4-6 page summaries of each chapter in addition to hundreds of examples worked out in detail, makes the text well suited to self-paced, individualized study. Review sections; detailed answers to odd numbered problems. LCL

APPLICATIONS, T, P*, L**. *Case Studies in Applied Mathematics.* Ed: Maynard Thompson. CUPM, 1976, 444 pp, (P). Nine modules on diverse applications intended for undergraduate instructors to enable them to introduce their students to the complete process of applying mathematics. Topics include voting systems, municipal street sweeping, steam generator flow, heat transfer in soil, several-species ecosystems. Two general essays introduce the volume with discussion of the process of applying mathematics and case studies of the trial teaching of these modules. LAS

APPLICATIONS, P. *Lecture Notes in Mathematics-525: Structural Stability, The Theory of Catastrophes, and Applications in the Sciences.* Ed: P. Hilton. Springer-Verlag, 1976, vi + 408 pp, $14.35 (P). Proceedings of an April 1975 conference at the Battelle Seattle Research Center; includes D.J.A. Trotman's revised version of E.C. Zeeman's important yet unavailable paper providing a complete proof of Thom's classification theorem for the elementary catastrophes. LAS

APPLICATIONS (PHYSICS), T(18), P. *Thermodynamics of Nonequilibrium Processes.* S. Wiśniewski, B. Staniszewski, R. Szymanik. Reidel, 1976, xiii + 274 pp, $36. Translated from the 1973 Polish original, this book presents "the fundamentals of nonequilibrium thermodynamics and its principal engineering applications." Prerequisites: classical thermodynamics and tensor calculus. Examples with solutions in place of exercises. JAS

Reviewers Whose Initials Appear Above

David F. Appleyard, Carleton; Ceceila Bleecker, Carleton; Paul J. Campbell, St. Olaf; John Dyer-Bennet, Carleton; Steven Galovich, Carleton; Loren Haskins, Carleton; Ih-Ching Hsu, St. Olaf; Paul S. Jorgensen, Carleton; Lorraine L. Keller, St. Olaf; Loren C. Larson, St. Olaf; R.W. Nau, Carleton; J. Arthur Seebach, St. Olaf; Lynn A. Steen, St. Olaf.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, NW, Washington, D. C. 20036. Items must be submitted at least five months before publication can take place.

PERSONAL ITEMS

Brock University: Associate Professor H. E. Bell has been promoted to Professor; Assistant Professor J. W. Auer has been promoted to Associate Professor.

Longwood College: Dr. W. L. Hightower has been appointed Assistant Professor; Assistant Professor R. S. Wu has been promoted to Associate Professor.

Mississippi University for Women: Dr. Larry Hoehn, University of Tennessee, has been appointed Assistant Professor; Professor Z. T. Gallion retired on May 31, 1976.

Ripon College: Dr. N. J. Loomer has been promoted from Instructor to Assistant Professor; Associate Professor C. W. Larson, Chairman of the Mathematics Department, has been promoted to Professor.

Shippensburg State College: Drs. J. G. Johnson and J. W. Crawley have been appointed Assistant Professors; Associate Professor J. S. Mowbray, Jr., has been promoted to Professor.

State University of New York College at Brockport: Assistant Professor D. S. Martin has been promoted to Associate Professor; Associate Professor Sanford Miller has been promoted to Professor; Professor Aziz Ibrahim has been awarded the Chancellor's Award for Excellence in Teaching.

The Citadel: Dr. J. I. Moore has been appointed Assistant Professor; Dr. J. L. Johnson has been appointed Instructor; Associate Professor and Acting Head of the Department of Mathematics I. S. Metts, Jr., has been appointed Head of the Department of Mathematics; Professor Lee P. Hutchison retired on July 1, 1976, with the title of Professor Emeritus.

The University of Toledo: Dr. C. B. Davis, University of Mew Mexico, has been appointed Assistant Professor; Associate Professor Paul Shields has been promoted to Professor.

Valparaiso University: Sister Mary Treanor has been appointed part-time Instructor; Assistant Professors Ruth K. Deters and Malcolm Reynolds have been promoted to Associate Professors; Professor M. G. Mundt has been appointed Chairman of the Mathematics Department.

York University: Assistant Professors Julia N. Brown and E. J. Mayland, Jr., have been promoted to Associate Professors.

Assistant Professor J. J. Avioli, Christopher Newport College, has been promoted to Associate Professor.

Mrs. Mary Jane Kohlenberg, Northeast Missouri State University, has been promoted from Instructor to Assistant Professor.

Professor George Ledin, Jr., University of San Francisco, has been appointed Chairman of the Department of Computer Science.

Mrs. Marsha Meredith, Hanover College, has been appointed Visiting Lecturer at Indiana University Southeast.

Associate Professor James Mettler, Pennsylvania State University, Schuylkill Campus, retired on July 1, 1976.

Associate Professor J. E. Morrill, DePauw University, has been promoted to Professor.

Professor M. F. Neuts, Purdue University, has been appointed to a Unidel Professorship in Statistics and Computer Science at the University of Delaware.

Associate Professor J. A. Pfaltzgraff, University of North Carolina, Chapel Hill, has been promoted to Professor.

Associate Professor K. L. Whipkey, Westminster College, has been promoted to Professor.

Mr. Frank Boehm, Tucson, Arizona, died on September 19, 1975. He was a member of the Association for thirty-eight years.

Dr. Anna M. Howe, DeWitt, New York, died on August 8, 1976, at the age of 93. She was a Charter Member of the Association.

Professor Mary Elizabeth Williams, Skidmore College, died on July 31, 1976, at the age of 66. She was a member of the Association for thirty-seven years.

MASS MEDIA INTERN PROGRAM

The American Association for the Advancement of Science announces a program to support up to 18 advanced social and natural science students as intern reporters, researchers, and production assistants in a variety of media for the summer of 1977. Interested students should write: Coordinator, Mass Media Intern Program, AAAS, 1776 Massachusetts Ave., N. W., Washington, D. C. 20036.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

CALENDAR OF FUTURE MEETINGS

Fifty-seventh Summer Meeting, University of Washington, August 14–16, 1977.

Sixty-first Annual Meeting, Atlanta, Georgia, January 6–8, 1978.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN, St. Francis College, Loretto, Pennsylvania, April 22–23, 1977.

FLORIDA, University of South Florida, Tampa, March 4–5, 1977.

ILLINOIS, Chicago Loop College, Chicago, May 6–7, 1977.

INDIANA, Wabash College, Crawfordsville, April 30, 1977.

INTERMOUNTAIN

IOWA, Drake University, Des Moines, April 22–23, 1977.

KANSAS, Tabor College, Hillsboro, April 2, 1977.

KENTUCKY, early April. Deadline for papers 6 wks. bef. mtg.

LOUISIANA-MISSISSIPPI, Friday-Saturday before February 20. Deadline for papers 3 mths. bef. mtg.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA, Saturday before Thanksgiving and last Saturday in April.

METROPOLITAN NEW YORK, Sarah Lawrence College, April 24, 1977.

MICHIGAN, Eastern Michigan University, Ypsilanti, May 6–7, 1977.

MISSOURI, University of Missouri, St. Louis, April 29–30, 1977.

NEBRASKA, Nebraska Wesleyan University, Lincoln, April 15–16, 1977.

NEW JERSEY, early November and early May.

NORTH CENTRAL, North Hennepin Community College, Minneapolis, Minnesota, April 29–30, 1977.

NORTHEASTERN, Saturday after Thanksgiving, and third week in June in odd-numbered years.

NORTHERN CALIFORNIA, first or second Saturday in February.

OHIO, Denison University, Granville, April 15–16, 1977.

OKLAHOMA-ARKANSAS, Oral Roberts University, Tulsa, Oklahoma, April 1–2, 1977.

PACIFIC NORTHWEST, second Saturday in June. Deadline for papers 6 wks. bef. mtg.

PHILADELPHIA, Moravian College, Bethlehem, November 19, 1977.

ROCKY MOUNTAIN, Metropolitan State College, Denver, Colorado, April 29–30, 1977.

SEAWAY, State University College at Buffalo, May 6–7, 1977.

SOUTHEASTERN, University of Alabama, Huntsville, April 1–2, 1977.

SOUTHERN CALIFORNIA, Loyola Marymount University, Los Angeles, March 12, 1977.

SOUTHWESTERN, Phoenix College, Phoenix, Arizona, April 22–23, 1977.

TEXAS, Baylor University, Waco, April 1–2, 1977.

WISCONSIN, University of Wisconsin, Oshkosh, April 29–30, 1977.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Washington, February 12–18, 1978.

AMERICAN MATHEMATICAL ASSOCIATION OF TWO YEAR COLLEGES

AMERICAN MATHEMATICAL SOCIETY, University of Washington, August 15–18, 1977.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION, University of North Dakota, Grand Forks, June 13–16, 1977.

ASSOCIATION FOR COMPUTING MACHINERY, Olympic Hotel, Seattle, Washington, October 17–19, 1977.

ASSOCIATION FOR SYMBOLIC LOGIC, Chicago Sheraton, Chicago, April 28–29, 1977.

ASSOCIATION FOR WOMEN IN MATHEMATICS

CANADIAN SOCIETY FOR HISTORY AND PHILOSOPHY OF MATHEMATICS/SOCIÉTÉ CANADIENNE D'HISTOIRE ET DE PHILOSOPHIE DES MATHÉMATIQUES, Hamilton, Ontario, June 2, 1977.

FIBONACCI ASSOCIATION

INSTITUTE OF MATHEMATICAL STATISTICS, Seattle, Washington, August 14–18, 1977.

MU ALPHA THETA

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Cincinnati, Ohio, April 20–23, 1977.

OPERATIONS RESEARCH SOCIETY OF AMERICA, San Francisco Hilton, May 9–11, 1977.

PI MU EPSILON

SCHOOL SCIENCE AND MATHEMATICS ASSOCIATION

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS

THE CARUS MATHEMATICAL MONOGRAPHS

The Monographs are a series of expository books intended to make topics in pure and applied mathematics accessible to teachers and students of mathematics and also to non-specialists and scientific workers in other fields.

These numbers are currently available:

1. *Calculus of Variations,* by G. A. Bliss.
2. *Analytic Functions of a Complex Variable,* by D. R. Curtiss.
3. *Mathematical Statistics,* by H. L. Rietz.
4. *Projective Geometry,* by J. W. Young.
6. *Fourier Series and Orthogonal Polynomials,* by Dunham Jackson.
7. *Vectors and Matrices,* by C. C. MacDuffee.
8. *Rings and Ideals,* by N. H. McCoy.
9. *The Theory of Algebraic Numbers* (Second edition), by Harry Pollard and Harold G. Diamond.
10. *The Arithmetic Theory of Quadratic Forms,* by B. W. Jones.
11. *Irrational Numbers,* by Ivan Niven.
12. *Statistical Independence in Probability, Analysis and Number Theory,* by Mark Kac.
13. *A Primer of Real Functions* (Second edition), by Ralph P. Boas, Jr.
14. *Combinatorial Mathematics,* by H. J. Ryser.
15. *Noncommutative Rings,* by I. N. Herstein.
16. *Dedekind Sums,* by Hans Rademacher and Emil Grosswald.
17. *The Schwarz Function and its Applications,* by Philip J. Davis.
18. *Celestial Mechanics,* by Harry Pollard.

One copy of each Carus Monograph may be purchased by individual members of the Association for \$5.00 each; additional copies and copies for nonmembers are priced at \$10.00 each. **Effective April 1977, prices will be \$6.50 and \$11.00, respectively.** (Orders for under \$10.00 must be accompanied by payment. Prepaid orders will be delivered postage and handling free.)

Orders should be sent to:

MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.
Washington, D.C. 20036

For Your Consideration...

The premise.

Calculus is not an esoteric subject.

The execution.

To teach calculus by making full use of students' native geometric understanding: Shenk introduces basic topics intuitively with a minimum number of examples, moving quickly on to the problem sets where outlines of solutions to the more difficult problems serve as further examples **at the point of need**. Pervading the problem sets are applied narrative problems of a range and inventiveness not found in existing texts. Computer-drawn sketches of curves and surfaces, particularly in the three-dimensional, help students visualize what they cannot draw. Rigorous treatments, such as the $\epsilon\,\delta -$ definition of the limit, are confined to optional appendices. Other items that may be included at the option of the instructor are historical notes on such topics as infinitesimals and exercises that make use of an electronic calculator or computer to help teach the limit processes.

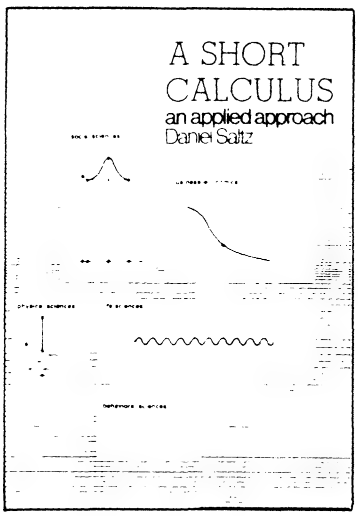
The result.

Calculus is made accessible to all.

The Contents.

Coordinate Planes and Graphs of Equations/Functions and Limits/The Derivative/Applications of Differentiation/The Integral/Applications of Integration/Transcendental Exercises/Techniques of Intergration/Separable Differential Equations/Conic Sections and Polar Coordinates/Curves and Vectors in the Plane/Vectors in Space and Solid Analytic Geometry/Partial Derivatives/Extremum Problems in Two and Three Variables/Multiple Integrals/Vector Analysis/Taylor's Theorem, L'Hospital's Rule, and Improper Integrals/Infinite Sequences and Series/Further Topics in Differential Equations/Tables/Index
Available Now! 900 pp., Clothbound.

Two New Calculus Texts...



New 2nd edition!

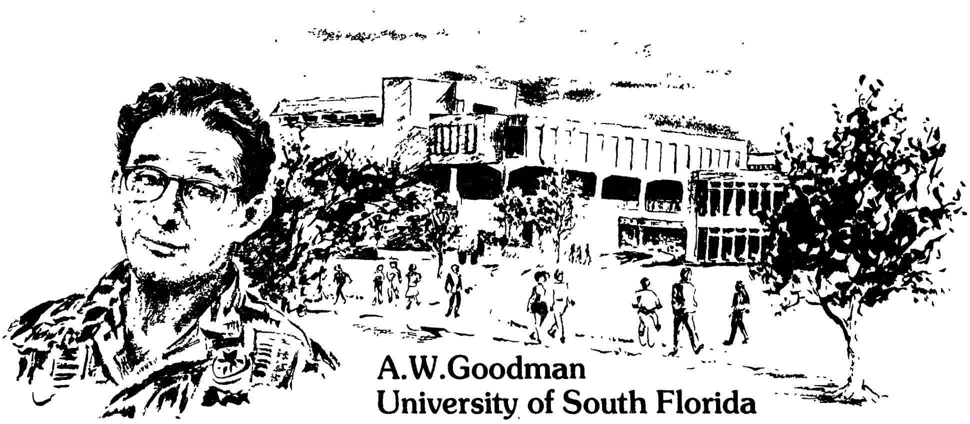
The second edition of this bestselling, geometrically oriented calculus text for non-math majors features nearly 100 applications—in worked-out examples, problem sets, and entire sections. Applications include 30 from biological and life sciences, 37 from business and economics, and many from psychology and other disciplines. Each section dealing with applications has been reworked for clarity and over 20 entirely new applications have been added. There are many new problems as well, and the problem sets are now graded to proceed from the routine to the most challenging. The text readily adapts to quarter, semester, or year-long calculus courses. **Instructor's Manual** includes solutions to even-numbered problems (odd-numbered solutions appear at the back of the text).

Applications for Biology-Medicine Include:

Differential Growth, the simple allometry formula as related to graphs/Temperature jumps on skin tissue/Animal growth (comparing tail length to body length of species of snake)/Functions comparing drugs to patients using drugs/Population of bacteria subject to antibacterial agent; population size of bacteria culture; weight of an organism/Rate of change of population of bacteria/Relationship between concentration of injected adrenaline to muscle response, blood pressure (stroke volume versus pressure in aorta during systole)/Maximum radius of trachea during coughing/Applications of max-min (sensitivity, i.e., reaction of body to injected drugs; animal diets and additives, photosynthesis, Ohm's law)/Rate of growth of size of bacteria as example of antiderivatives/Sequential model for population size/Population growth and drug dissipation as related to geometric sequences/Drug injection as related to a geometric sum/Cell division (as introductory example for exponential function)/Concentration of drug injected into blood (exponential model)/Exponential sequence for population/Relationship between surface area of human body and weight/Exponential population growth — and many, many, more from business, economics, social science and psychology... Available January 1977, 448 pp., Clothbound.

from Goodyear Publishing.

For your examination copies, write to me at the address below, telling me the name of the course for which the book is to be considered, the probable enrollment and the text you are presently using. Write to James Van/ Dept AMM/ Goodyear Publishing/ Box 2113/ Santa Monica, CA 90272.



A.W.Goodman
University of South Florida

CALCULUS FOR THE SOCIAL SCIENCES

Table of Contents

For further information, write: **Textbook Marketing Division**

W. B. Saunders Company
West Washington Square, Philadelphia, Pa. 19105

THE MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

AN OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 84

NUMBER 4

CODEN: AMMYAE

CONTENTS

(*Continued on inside cover*)

APRIL 1977

NOTICE TO AUTHORS

Research papers per se are unsuitable; see statement of policy (Vol. 81, p. 1).

Please follow the format in current issues of the MONTHLY. Manuscripts must be legibly typewritten or reproduced from typewritten copy, double spaced with wide margins. Three copies should be submitted to the appropriate editor and one kept by the author as protection against loss. The author's full address *must* appear at the end of the manuscript.

Backlog: Main Articles 15 months, Math. Notes 17 months, Research Problems 14 months, Classroom Notes 14 months, Math. Education 16 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to R. P. BOAS, Department of Mathematics, Northwestern University, Evanston, IL 60201; NOTES, etc.: to the corresponding Associate Editor; REPRINT PERMISSION: to LEONARD GILLMAN, Mathematical Association of America, University of Texas, Austin, Texas 78712 (see also the copyright notice below); ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N.Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N. W., Washington, D.C. 20036; BACK ISSUES: Contact P. and H. Bliss Co., Middletown, CT 06457.

Annual dues for members of the Association, including a subscription to the American Mathematical Monthly, are $18.00 for each of the first two years of membership and $22.00 thereafter. Student Membership is available with annual dues of $12.00. For nonmembers the subscription price is $25.00.

PUBLISHED BY THE ASSOCIATION at Washington, D.C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D.C., and additional mailing offices.

PARTIAL SUMS OF INFINITE SERIES, AND HOW THEY GROW

R. P. BOAS, JR.

Introduction. The four chapters of this article are separate informal essays on topics connected with the problem of finding out, as accurately as possible, how fast a convergent series converges and how fast a divergent series diverges. Chapter I is a general discussion with numerical examples. Chapter II is a self-contained development of the Euler–Maclaurin formula, the formula that lets us approximate the partial sums of series (if the series have sufficiently simple structure) by integrals, with errors that can be estimated. The Euler–Maclaurin formula, in its more sophisticated versions, involves the Bernoulli numbers, and Chapter III develops the properties of these numbers to the point where some of their interesting applications can be made. In Chapter IV I show how sums of some convergent series, the Euler constants for some divergent series, and good approximate formulas for partial sums can be actually calculated numerically.

I hope that some of this material, which is really quite elementary, may find its way into the undergraduate curriculum, perhaps in place of some of the dull and not particularly useful material that is traditionally there.

I am much indebted to John W. Wrench, Jr., who provided me with some essential numerical data, as well as advice in my first ventures into numerical analysis; and to Leonard Evens, who has helped me learn to talk to the computer without its scolding me too often. I have used portions of these essays as talks at section meetings of the Association and elsewhere, as well as in my retiring Presidential Address to the Association, and I have profited from useful comments made by members of my audiences.

I. Convergence and divergence. I think that mathematics courses usually give students misleading ideas about convergence and divergence of series. Textbooks spend a lot of time on tests for convergence that are of little practical value, since a convergent series either converges rapidly, in which case almost any test will do; or it converges slowly, in which case it is not going to be of much use unless there is some way to get at its sum without adding up an unreasonable number of terms. Of course the definition of "reasonable" changes with the technology. At present it is not too hard to add up a few million terms, at least on some computers; I have access to one that can theoretically add on a term in about 10 microseconds, so that it should be able to add a million terms in 10 seconds (actually it seems to take something between 10 and 30 times as long, since (among other things) it has to compute the terms in order to add them, but a few hundred seconds is not too unreasonable); some computers are about ten times as fast. However, 10^{12} terms seem to be beyond the range of any current computer and 10^{15} beyond any conceivably practical one. Indeed, even if we optimistically count on a term per microsecond, 10^{12} terms would take about $11\frac{1}{2}$ days, 10^{15} terms would take about 32 years, and 10^{23} terms would take the estimated age of the universe. It can be argued that the fastest imaginable computer would add a term while light crosses a (classical) electron (about 10^{-23} sec); such a computer, started off at the beginning of the universe, would have added only about 10^{40} terms by now. However, a computer can sometimes be used to get indirectly at the sum of a series when the sum cannot be found by brute force addition; we shall see some examples later.

Our ideas about convergence seem mostly to be formed on a basis of power series, which typically converge rather fast. Consider, for example, a geometric series with a ratio close to 1; we are fortunate enough to have a formula for its partial sums,

$$\sum_{k=0}^{N} x^k = \frac{1-x^{N+1}}{1-x} = \frac{1}{1-x} - \frac{x^{N+1}}{1-x}.$$

Thus if we want to compute the sum by adding up the terms from the beginning until we have an error

of less than ε, we need $N+1$ terms where N is large enough to make $x^{N+1} < \varepsilon(1-x)$ (assuming $x > 0$). We can easily calculate that when $\varepsilon = 0.005$ (this is usually considered to be two-decimal-place accuracy, but see the appendix to this chapter), and x is even as large as 0.95, we still need only 162 terms. If we want much greater accuracy, say $\varepsilon = 5 \times 10^{-100}$, we will need 4517 terms, which would be tedious by hand but trivial for a high-speed computer.

However, as soon as we get away from power series the situation deteriorates. For example, if we want to add terms of $\Sigma_1^\infty 1/n^2$ until the error is less than 0.005, we need (precisely) 200 terms (as will be verified later), more than for $\Sigma(0.95)^n$; and each additional decimal place demands 10 times as many terms as the previous one. Thus to get 10 decimal places we should have to add up 2×10^{10} terms, which is rather impractical. It happens that we know, in fact, that $\Sigma_{n=1}^\infty 1/n^2 = \pi^2/6$, and so we can get the sum very accurately if we have a sufficiently good value for π (10^6 decimal places have been calculated, although only about 10^5 have been published, as far as I know). However, we can make similar statements about rates of convergence even when we do not know the sum. One of the objects of this chapter is to show how such computations can be made. It is a different, but related, and somewhat more difficult, problem to calculate the sum of a series when it would take an unreasonable or impossible number of terms to get it directly to a desired degree of accuracy. For example, $\Sigma_{n=2}^\infty n^{-1}(\log n)^{-2}$ would require about 10^{87} terms (the exact number is given below, on p. 240) to get its sum to 2 decimal places, but the sum is known, by indirect methods, to be approximately 2.10974. (See Chapter 4.) For $\Sigma_{n=3}^\infty 1/(n \log n\, (\log\log n)^2)$, the number of terms required for 2-decimal-place accuracy cannot be written out in full, since it has more digits that there are particles in the universe.

The situation for divergence is similar. Except possibly at the endpoints of the interval of convergence, a divergent power series has terms that become infinite, so that its divergence is rather intuitive. On the other hand, a series like the harmonic series $\Sigma\, 1/n$, which both has terms that approach zero and diverges slowly, can be completely misleading to someone who tries to compute partial sums without thinking carefully about what is happening. For example, after 83 terms the partial sums of the harmonic series exceed 5, and each term is adding only a little more than 0.01, so that one might be tempted to assume that the series is going to converge to something not much greater than 5. However, after 12367 terms the sum has reached 10, and even though the individual terms are now less than 10^{-4}, one might change one's mind. After a few more than 6 trillion terms the sum has reached 30, and after something more than 1.5×10^{43} terms it reaches 100 (see **[5]** for details). (In practice, the computer would long since have rounded off the terms to 0, and nothing would be happening.) Of course direct computation of these large numbers of terms is out of the question. They can be got at by the so-called Euler–Maclaurin formula (derived in Chapter 2), if we are studying series of the kind that appear in textbooks as illustrations of the integral test. These are series $\Sigma f(n)$, where f is positive and decreasing, $|f'|$ also decreases, and f is the derivative of a function whose values we can find easily. We can use one of the simplest versions of the Euler–Maclaurin formula; usually the so-called "second form" is somewhat more convenient. For a convergent series, the Euler–Maclaurin formula produces the following expression for the remainders of the series:

$$R_n \equiv \sum_{k=n+1}^{\infty} f(k) = \int_{n+1/2}^{\infty} f(t)dt + \int_{n+1/2}^{\infty} P_1(t)f'(t)dt,$$

where $P_1(t) = t - [t] - 1/2$ ($[t]$ being the integral part of t). Since P_1 has period 1 and average value 0 over a period, is odd around $n + 1/2$, and positive between $n + 1/2$ and $n + 1$, we have

$$\int_{n+1/2}^{\infty} P_1(t)f'(t)dt < 0,$$

as is clear from a sketch (remember that f' is negative and increasing). Also, since f' is negative and $|f'|$ decreases, we can apply the second mean-value theorem (see, for example, **[22]**, p. 163) and get

$$R_n = \int_{n+1/2}^{\infty} f(t)dt + f'(n+\tfrac{1}{2})\int_{n+1/2}^{x} P_1(t)dt$$

for some $x > n + 1/2$; it is again clear from a sketch that the last integral is between 0 and 1/8. Consequently the remainder R_n satisfies

$$\int_{n+1/2}^{\infty} f(t)dt > R_n > \int_{n+1/2}^{\infty} f(t)dt + \tfrac{1}{8}f'(n+\tfrac{1}{2}).$$

If $y = \psi(x)$ is the inverse of $x = \int_y^\infty f(t)dt$, the preceding inequality says that R_n is nearly equal to some specified $\varepsilon > 0$ when $n + 1/2$ is nearly equal to $\psi(\varepsilon)$.

More precisely, we shall surely have $R_n < \varepsilon$ together with $R_{n-1} > \varepsilon$ (so that precisely n terms of $\Sigma_1^\infty f(k)$ are required to get a remainder less than ε) provided that we have both

$$\int_{n+1/2}^{\infty} f(t)dt < \varepsilon \quad \text{and} \quad \int_{n-1/2}^{\infty} f(t)dt + \tfrac{1}{8}f'(n-\tfrac{1}{2}) > \varepsilon.$$

As illustrations, let us verify some statements made earlier.

(a) $f(k) = 1/k^2$. Then

$$\int_{n-1/2}^{\infty} f(t)dt = \frac{2}{2n-1},$$

and we have remainder less than ε for the first time at n, provided that

$$\frac{2}{2n+1} < \varepsilon \quad \text{and} \quad \frac{2}{2n-1} - \frac{2}{(2n-1)^3} > \varepsilon.$$

If $\varepsilon = 1/200 = 0.005$, the inequality

$$\frac{2}{2n+1} < \frac{1}{200}$$

says that $2n + 1 > 400$, $n > 199.5$, in other words $n \geqq 200$; and with $n = 200$ we have

$$\frac{2}{2n-1} - \frac{2}{(2n-1)^3} > \frac{1}{200}$$

with a good deal to spare. Similarly, if $\varepsilon = 1/2000$, $n = 2000$, and so on.

This calculation incidentally illustrates the fallacy of the advice that is sometimes given, namely that in calculating the sum of a series one can stop when the terms become smaller than the admissible error. Here (for example) the 200th term is less than 4×10^{-5} but the error in the sum at this point is much larger, in fact close to 5×10^{-3}.

(b) $f(x) = x^{-1}(\log x)^{-2}$. We have

$$\int_x^{\infty} f(t)dt = 1/\log x,$$

and so we shall have $R_n < 1/200$ for the first time, provided that

$$\frac{1}{\log(n+\frac{1}{2})} < \frac{1}{200} \quad \text{and} \quad \frac{1}{\log(n-\frac{1}{2})} - \frac{1}{8(n-\frac{1}{2})[\log(n-\frac{1}{2})]^2} > \frac{1}{200}.$$

The first inequality requires $n + \frac{1}{2} > e^{200}$ = approximately 7.2×10^{86} and hence suggests that n is the first integer exceeding $e^{200} - \frac{1}{2}$. The second inequality can in fact be shown to hold for this n, and so

(since e^{200} is known to sufficient accuracy [**6**]) we have

$$n = 7.22597\ 37681\ 25749\ 25817\ 74770\ 42189\ 30569\ 73568\ 74428\ 52731\ 92840\ 32697\ 89123\ 22190\ 93614\ 73891\ 66156\ 2 \times 10^{86}.$$

The actual number of terms required is 1 less than this since our series started at $n = 2$.

The nth term is approximately 6.9×10^{-87}, so that the cautionary remark at the end of (a) is even more cogent here.

The preceding calculations were rather simple because we were able to evaluate $\varphi(y) = \int_y^\infty f(t)dt$, and its inverse $\psi(x)$, in terms of tabulated, or at least easily computed, functions. However, there are serious difficulties if we go beyond series with the simplest kinds of terms. Let us consider, for example, the series $\Sigma_{n=2}^\infty n^{-2} \log n$ and $\Sigma_{n=2}^\infty 1/(n^2 \log n)$; we can expect the first to converge somewhat more slowly than $\Sigma 1/n^2$ and the second to converge somewhat faster; but the difference in rapidity ought not to be very great, since the logarithm increases so slowly (remember that since log 10 is about 2.3, the logarithm of the number of particles in the universe, supposed to be about 10^{80}, is less than 200). The inequalities that have to be satisfied to make the remainders less than ε are, respectively,

$$\int_{n+1/2}^\infty \frac{\log x}{x^2} dx < \varepsilon \quad \text{and} \quad \int_{n+1/2}^\infty \frac{dx}{x^2 \log x} < \varepsilon.$$

The first of these integrals is easy to evaluate:

$$\int_{n+1/2}^\infty \frac{\log x}{x^2} dx = \frac{1 + \log(n + \frac{1}{2})}{n + \frac{1}{2}}.$$

However, how do we find n from

$$\frac{1 + \log(n + \frac{1}{2})}{n + \frac{1}{2}} < \varepsilon\,?$$

For $\varepsilon = \frac{1}{2} \times 10^{-2}$ we can find $n = 1686$ by trial and error, so that 1685 terms are needed (we started the series at $n = 2$). For $\varepsilon = \frac{1}{2} \times 10^{-10}$, this approach seems impractical. It is slightly easier to try to find the first n such that

$$\frac{n + \frac{1}{2}}{1 + \log(n + \frac{1}{2})} > 1/\varepsilon.$$

Fortunately the technique for approximating inverse functions, barely mentioned in [**13**] (p. 42), has recently been developed to the point where it can be applied to this problem [**8a**], [**3a**]. The successive approximations to the inverse of $y = x/\log x$ are (if we write Ly for $\log y$, L_2y for $\log\log y$)

$$(*) \qquad x = y\{Ly + L_2y\} + \frac{L_2y}{Ly} + \frac{yL_2y}{(Ly)^2}\{1 - \tfrac{1}{2}L_2y\} + \frac{yL_2y}{(Ly)^3}\{1 - \tfrac{3}{2}L_2y + \tfrac{1}{3}(L_2y)^2\}$$
$$+ \frac{yL_2y}{(Ly)^4}\{1 - 3L_2y + \tfrac{11}{6}(L_2y)^2 - \tfrac{1}{4}(L_2y)^3\} + \cdots.$$

To invert $y = x/(1 + \log x)$, write $u = xe$, $v = ye$, so that $v = u/\log u$; substituting from (*) and expanding, we get

$$x = y\left\{1 + \log y + \log(1 + \log y) + \frac{\log(1 + \log y)}{1 + \log y}\right\} + \cdots$$

(where each Ly in (*) has been replaced by $1 + \log y$ and each L_2y by $\log(1 + \log y)$). If we write A_n for the approximation obtained by going through terms in $(1 + \log y)^{-n}$, we get the following

successive approximations when $y = 200$ (for the index n at which the error becomes less than 0.005):

$$A_0 = 1627.7, \qquad n = 1628$$

$$A_1 = 1686.16, \qquad n = 1687$$

$$A_2 = 1686.90, \qquad n = 1687$$

$$A_3 = 1685.97, \qquad n = 1686$$

$$A_4 = 1685.998, \qquad n = 1686.$$

We see that A_0 is quite a poor approximation; A_1, the approximation given in **[13]**, is not quite good enough for our purposes, and in fact it is only with A_3 (or the truncated version of A_3 given in **[3a]**, p. 95) that we get the correct value of n. For the other entries for this series we can get the numbers of significant figures claimed by simply using A_2. It is clear that the series does converge more slowly than $\Sigma\, 1/n^2$.

For $\Sigma\, 1/(n^2 \log n)$ the situation is worse. We have to have $E_1(\log(n + 1/2)) < \varepsilon$, where $E_1(x)$ is the exponential integral (also known as $-\mathrm{Ei}(-x)$),

$$E_1(x) = \int_x^\infty t^{-1} e^{-t} dt.$$

The most extensive tables I know of **[18]** show that $E_1(\log x)$ becomes less than $1/2 \times 10^{-2}$ when x is about 3.77, which suggests that $E_1(\log(n + 1/2)) < 1/2 \times 10^{-2}$ when $n = 43$ (42 terms), and the correctness of this number of terms can be confirmed by actually adding them up. The tables go far enough to show that n is about 10^9 for $\varepsilon = 1/2 \times 10^{-10}$. Getting a useful formula for the inverse of $E_1(x)$ seems to be out of the question. Very roughly we can say that $E_1(x)$ is asymptotically e^{-x}/x, so that $E_1(\log n)$ is asymptotically $1/(n \log n)$, and $E_1(\log n)$ becomes less than ε at about $n = (1/\varepsilon)/\log(1/\varepsilon)$.

An instructive example of a very slowly convergent series is the series obtained by deleting from the harmonic series the reciprocals of the integers whose decimal representations contain at least one zero. Contrary to what intuition might suggest, it converges; but it converges so slowly that even after more than 10^{18} terms the remainder still exceeds 1. (Consequently there is no possibility of adding up the terms until the remainder is negligible.) Although the series is too irregular for the Euler–Maclaurin formula to be useful, its sum can be calculated to several decimal places by getting both upper and lower bounds for the (rather large) remainder. For example, after 66429 terms (those involving integers with 5 or fewer digits) the partial sum is about 9.832128 and the remainder exceeds 13; the sum correct to 6 decimal places is 23.103448 (it has been calculated to 20 decimal places by R. Baillie).

Table 1 on page 242, adapted from **[13]**, illustrates the rate of convergence of various series. In fact only series (1), (2), and (3) can be handled satisfactorily by the Euler–Maclaurin formula. For (4), the remainders can of course be calculated exactly; and series (5) through (8) converge so fast that the remainders are not difficult to estimate. Note that for a series like (3) ($s = 100$) a few terms give quite high accuracy, but for very high accuracy a geometric series like (4) ($x = 0.9$), although it starts off less impressively, does better eventually.

We also want to know how many terms of a divergent series would have to be added up to get a specified sum. To get really precise results for a series $\Sigma_1^\infty f(k)$, we need to know the value of the Euler constant

$$\gamma = \lim_{n\to\infty} \left\{ \sum_1^n f(k) - \int_1^n f(t)dt \right\}$$

for the series. (The original Euler constant belongs to $\Sigma\, 1/n$.) Chapter IV explains how Euler constants

TABLE 1

$$(1)\ \sum_3^\infty \frac{1}{n\log n\,(\log\log n)^2}\quad (2)\ \sum_2^\infty \frac{1}{n\,(\log n)^2}\quad (2\text{a})\ \sum_2^\infty \frac{\log n}{n^2}\quad (3)\ \sum_1^\infty \frac{1}{n^s}\quad (3\text{a})\ \sum_2^\infty \frac{1}{n^2\log n}$$

$$(4)\ \sum_0^\infty x^n\quad (5)\ \sum_1^\infty \frac{1}{n!}\quad (6)\ \sum_1^\infty \frac{1}{n^n}\quad (7)\ \sum_1^\infty x^{n^2}\quad (8)\ \sum_1^\infty n^{-n^n}$$

Series	Sum	Number of terms required to calculate the sum with error less than 1/2 times			
		10^{-2}	10^{-10}	10^{-100}	10^{-1000}
1	38.406768	$T(3.14\times 10^{86})$	$T_4(1)$	$T_4(100)$	$T_4(1000)$
2	2.109743	7.23×10^{86}	$T(8.7\times 10^9)$	$T_2(100)$	$T_2(1000)$
$3(s=1.1)$	10.584448	10^{33}	10^{113}	10^{1013}	10^{10013}
$3(s=1.5)$	2.612375	160000	1.6×10^{21}	1.6×10^{201}	1.6×10^{2001}
2a	0.937548	1685	5.47×10^{11}	4.7×10^{102}	4.6×10^{1003}
$3(s=2)$	1.644934	200	2×10^{10}	2×10^{100}	2×10^{1000}
3a	0.605522	42	10^9	10^{98}	10^{997}
$3(s=10)$	1.000995	1	11	1.1×10^{11}	1.1×10^{111}
$3(s=100)$	$1+7.9\times 10^{-31}$	1	1	10	1.21×10^{10}
$4(x=.9)$	10	73	247	2214	21883
$4(x=.5)$	2	9	36	335	3324
$4(x=.1)$	10/9	3	11	101	1001
5	1.718282	5	13	70	450
6	1.291286	3	10	57	386
$7(x=.9)$	2.230273	7	15	46	147
$7(x=.5)$	0.564468	2	5	18	57
$7(x=.1)$	0.001000	1	3	10	31
8	1.062500	2	2	3	4

Note: $T(x)=T_1(x)=10^x$, $T_n(x)=T(T_{n-1}(x))$.

can be calculated. Once again the calculations will go through fairly easily provided that $\int f(x)dx$ is a well-tabulated function.

Let n_A be the number of terms of $\Sigma_{k=1}^\infty f(k)$ we have to take to get a sum s_n that is at least equal to A for the first time. Even the simplest version of the Euler–Maclaurin formula will produce quite accurate results when A is large. If we write $\varphi(n)=\int_1^n f(t)dt$, the formula (see Chapter II) tells us that

$$s_n=\varphi(n)+\tfrac{1}{2}f(n)+\gamma-\int_n^\infty P_1(t)f'(t)dt.$$

If, as often happens, the remainder integral can be neglected, we see that s_n reaches A when $\varphi(n)+\frac{1}{2}f(n)+\gamma$ is close to A, or (since $f(n)$ itself is small for large n) when n is nearly $\varphi^{-1}(A-\gamma)$. For functions f which, with their derivatives, approach zero sufficiently fast at ∞ (roughly speaking, anything made out of powers of x and logarithms will do, as long as it approaches zero), more elaborate calculations [**4**] show that when A is an integer, n_A is either $m=[\varphi^{-1}(A-\gamma)]$ or $m+1$; the first or second case occurs according as

$$\varphi^{-1}(A-\gamma)<m+\tfrac{1}{2}+\tfrac{1}{24}\left(\frac{|f'(m)|}{f(m)}-\varepsilon\right)$$

or

$$\varphi^{-1}(A-\gamma)>m+\tfrac{1}{2}+\tfrac{1}{24}\left(\frac{|f'(m)|}{f(m)}+\varepsilon\right),$$

where ε is a positive function that approaches zero as $A \to \infty$. For any particular (not too complicated) f one can make even more precise statements. The one just made almost (but not quite) says that n_A is the integer closest to $\varphi^{-1}(A - \gamma)$ (when A is an integer); this may in fact be the actual value of n_A for large integral values of A (it need not be when A is not an integer); but this seems not to have been proved for any nontrivial case.

Table 2 on page 244 (adapted from [13]) shows how fast the partial sums of some divergent series grow. The series are arranged in order of increasingly slow divergence; it will be noticed that the first few terms can give a very poor idea of the rate of divergence, and that a logarithmic factor has considerable influence in spite of the slow rate of increase of the logarithm function. The series $\Sigma_1^\infty n^{-0.1}$ does not appear in Table 2. Since $n^{0.1}$ grows (eventually) faster than $\log n$, this series should diverge more slowly than $\Sigma\, 1/\log(n+1)$, but this would not show up in Table 2, since even at 10^6 the partial sums of $\Sigma n^{-0.1}$ are about 4.1×10^6. Series 2 does not catch up until somewhere around 10^{15}.

The smaller entries in Table 2 (those less than 10^6) were calculated by direct addition by machine and checked, whenever practicable, by evaluating $\varphi^{-1}(A - \gamma)$; the other entries were got directly from $\varphi^{-1}(A - \gamma)$. For all the series except (1) and (2), φ is expressible in closed form; for (2), φ has been sufficiently well tabulated; in order to get the last entry for (1) it was necessary to begin by tabulating the corresponding $\varphi(x)$.

Appendix. *The meaning of "two-decimal place accuracy."*

We have been calculating such things as the number of terms of a given convergent series that we would have to add up to get the sum with an error of less than 0.005. This is often taken as synonymous with "correct to two decimal places," but it is not always what seems to be intended by "computing to two decimal places." There is no confusion about what the correct two-decimal place value of $\sqrt{2}$ is: it is 1.41, since $\sqrt{2}$ has been calculated to more than enough accuracy for us to be sure that it correctly rounds to this value. But if we are calculating $\sqrt{2}$ by some algorithm, when are we entitled to stop? I presume that the answer is that we go on until what we have, rounded to 2 decimal places, is the correct value rounded to the same number of decimal places. If, for example, we are adding up a series of positive terms to get approximations to $\sqrt{2}$, and we get to, say, 1.406, which rounds to 1.41, we have done what was required, even though the difference between 1.406 and $\sqrt{2}$ is more than 0.005. On the other hand, if we had reached 1.4191 we would have an error less than 0.005 but this value rounds to 1.42, which is wrong.

The problem of computing the number of terms of $\Sigma 1/n^2$ that must be added to get the sum to 2 decimal places illustrates the point quite well. We have seen that precisely 200 terms are required to get the sum with an error less than 0.005. The exact value is 1.644934..., which rounds, to 2 decimal places, to 1.64, and we get this as soon as we have added enough terms to get anything that rounds to 1.64, i.e., anything greater than 1.635. We can calculate in the same way as before that this requires far fewer terms than before, namely 101.

On the other hand, an error of at most 0.0005 requires 2000 terms, but the correct 3-decimal value of the sum is 1.645, and to get this we must take enough terms to reach more than 1.6445, in fact 2304 terms (with sum 1.64450013).

Here the calculations were no more difficult than those required to get an error of at most 0.005 or 0.0005; but this is only because the sum is known already to more than the required number of decimal places. We can find the number of terms required to get at most a specified error without knowing the numerical value of the sum, but not the number of terms required to get a specified number of decimal places. To find the number of terms required to get 1000 decimal place accuracy in the strict sense requires knowing the sum to more than 1000 decimal places, and there are rather few series for which the sums have been calculated with this much precision.

TABLE 2

$$(1)\sum_1^\infty \frac{1}{\log\log(n+2)} \quad (2)\sum_1^\infty \frac{1}{\log(n+1)} \quad (3)\sum_1^\infty \frac{1}{\sqrt{n}} \quad (3a)\sum_1^\infty \frac{\log n}{n}$$

$$(4)\sum_1^\infty \frac{1}{n} \quad (5)\sum_1^\infty \frac{1}{(n+1)\log(n+1)} \quad (6)\sum_1^\infty \frac{1}{(n+2)\log(n+2)\log\log(n+2)}$$

Series	Value of γ	Number of terms required to make the sum greater than									
		3	4	5	6	7	10	20	100	1000	10^6
1	7.21848	1	1	1	1	1	1	6	112	1812	2.62×10^6
2	0.80193	3	5	7	9	12	20	56	489	7764	1.55×10^7
3	0.539645	5	7	10	14	18	33	115	2574	250731	2.50×10^{11}
3a	-0.0728158	12	17	24	33	43	89	565	1.39×10^6	2.65×10^{19}	1.53×10^{614}
4	0.577216	11	31	83	227	616	12367	2.72×10^8	1.5×10^{43}	1.1×10^{434}	$T(4.3\times 10^5)$
5	0.428166	8717	5.1×10^{10}	1.3×10^{29}	1.4×10^{79}	1.4×10^{215}	1.6×10^{4321}	$T(10^8)$	$T(5\times 10^{42})$	$T(4\times 10^{433})$	$T_2(4.3\times 10^5)$
6	2.299927	1	3	56	3.1×10^{19}	$T(1.3\times 10^4)$	$T(.7\times 10^{90})$	$T_2(2\times 10^6)$	$T_2(1.1\times 10^{41})$	$T_2(8\times 10^{431})$	$T_3(4.3\times 10^5)$

Note: $T_1(x)=T(x)=10^x$, $T_n(x)=T(T_{n-1}(x))$.

II. Comparing series with integrals.

1. You are probably willing to believe that there must be a close connection between the series $\Sigma f(n)$ and the integral $\int f(t)dt$. A picture makes it plausible (see Figure 1). Here the areas of the rectangles add up to $\Sigma_{n=a}^{b} f(n)$ and the area under the curve is $\int_a^b f(x)dx$; the difference looks pretty small in comparison with either. Of course I have drawn the picture in the most favorable case; a more "general" function might look more like Figure 2, and the comparison would not seem as convincing. One needs to suppose that the graph looks more like the first picture than like the second, perhaps that f is continuous and decreasing. Then one may ask, how close is the connection between the series and the integral? Also, suppose that we have an answer to this question; what can we do with it?

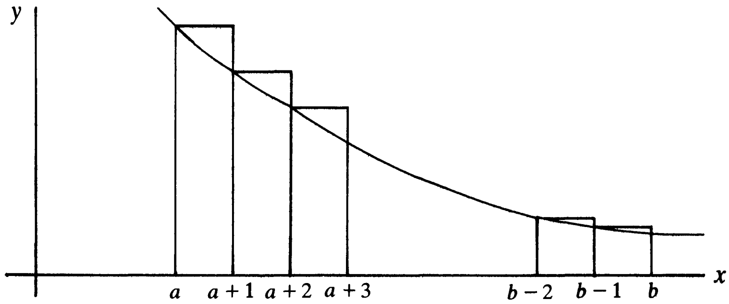


FIG. 1

A question that seems completely unrelated to this one is apt to occur to anyone who looks at a list of the power series of elementary functions. If you just write down the first few terms of the power series for, say, tan x, the coefficients seem completely irregular; but the tables usually give formulas for a number of such series in terms of some mysterious numbers B_n. What is going on here?

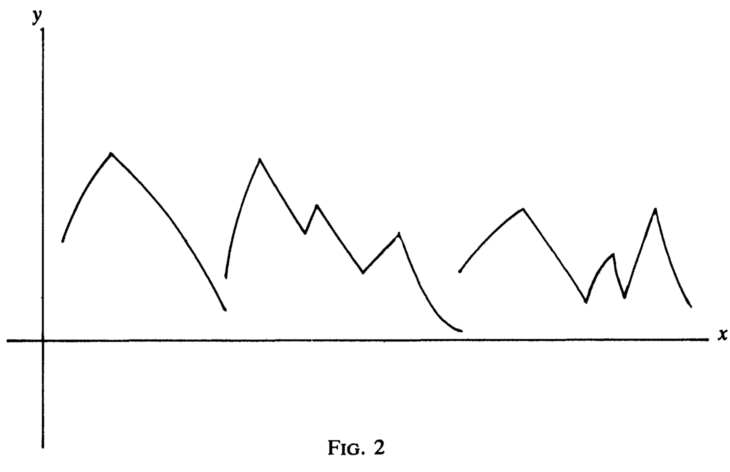


FIG. 2

Strangely enough, trying to answer the first question will lead us right to the answer to the second.

We shall suppose that f is continuous and has as many derivatives as the formulas require. In most applications f will also be monotonic, but this assumption is not needed to get the formulas, only in applying them. To keep the formulas simple we shall work with sums $\Sigma_{k=1}^{n} f(k)$, but we can replace the limits by m and $m+n-1$ whenever we like. (Technically speaking, $\Sigma_m^{m+n-1} f(k)$ is the same as $\Sigma_1^n f(m+k-1)$, so that we are really just applying the same formula to a different function.)

There is really no compelling reason for starting the series and the integral at the same place, as long as the distance between the limits of summation or integration is the same, so we shall try to compare $\Sigma_1^n f(k)$ with $\int_{1-a}^{n-a} f(t)dt$ with a between 0 and 1. In point of fact, the only useful cases seem to be $a = 0$ and $a = 1/2$, and about the only reason for using a general value of a is to get both formulas at once.

The motivation for the first step is that integrals are usually more manageable than sums, so we try to express the difference between $\Sigma f(k)$ and $\int f(t)dt$ as an integral.

The most straightforward way to do this is to use the notation of Stieltjes integrals. If you are not familiar with this, you can either stop and look it up (for example in [**21**] or [**22**]), or skip ahead to formula (1.4). Once you see the formula you can also see how to verify it directly; but it is harder to see how it could have been invented in the first place. (Actually the formula antedates Stieltjes by a good many years.)

We start by writing the sum as an integral,

$$\sum_{k=1}^{n} f(k) = \int_{1-a}^{n+1-a} f(t)d[t], \qquad 0 < a < 1, \tag{1.1}$$

where as usual $[t]$ means the greatest integer that does not exceed t. If $a = 0$, the integral in (1.1) is to be read as

$$\int_{1-}^{n+1-} f(t)d[t].$$

It will be convenient to modify (1.1) to read

$$\sum_{k=1}^{n} f(k) = \int_{1-a}^{n+1-a} f(t)d\{[t] + \tfrac{1}{2}\}.$$

Then we have

$$\sum_{k=1}^{n} f(k) - \int_{1-a}^{n+1-a} f(t)dt = \int_{1-a}^{n+1-a} f(t)d\{[t] + \tfrac{1}{2} - t\}.$$

The function in $\{\cdots\}$ will occur so frequently that we need a name for it; we shall write $t - [t] - \frac{1}{2} = P_1(t)$, so that

$$\sum_{k=1}^{n} f(k) - \int_{1-a}^{n+1-a} f(t)dt = -\int_{1-a}^{n+1-a} f(t)dP_1(t).$$

Notice that P_1 is a periodic function of period 1.

We now have the difference between sum and integral expressed as an integral. The immediate reaction of every analyst to an integral is to integrate it by parts, so we do this:

$$\sum_{k=1}^{n} f(k) - \int_{1-a}^{n+1-a} f(t)dt = f(1-a)P_1(1-a) - f(n+1-a)P_1(n+1-a) + \int_{1-a}^{n+1-a} P_1(t)df(t).$$

Since P_1 has period 1, we have $P_1(n+1-a) = P_1(1-a)$, and if we refer to the definition of P_1 we see that $P_1(1-a) = \frac{1}{2} - a$. Hence

$$\sum_{k=1}^{n} f(k) - \int_{1-a}^{n+1-a} f(t)dt = \{f(1-a) - f(n+1-a)\}(\tfrac{1}{2} - a) + \int_{1-a}^{n+1-a} P_1(t)df(t). \tag{1.3}$$

If f has a continuous derivative, this becomes

$$\sum_{k=1}^{n} f(k) - \int_{1-a}^{n+1-a} f(t)dt = (\tfrac{1}{2} - a)\{f(1-a) - f(n+1-a)\} + \int_{1-a}^{n+1-a} P_1(t)f'(t)dt. \tag{1.4}$$

This is our basic formula. If we want to start from (1.4) we can verify it without Stieltjes integration by writing the integral on the right as a sum of integrals over $(1-a,1)$, $(1,2),\ldots,(n,n+1-a)$, integrating by parts, and taking account of the discontinuities of P_1.

We now have a formula for the difference on the left of (1.4), and if we are lucky this ought to let us deduce something about how small that difference is. We can think of the right-hand side of (1.4) as a remainder term when the sum is replaced by the integral (or when the integral is replaced by the sum). It should be clear that the most promising situation is going to occur when $f'(t)$ is much smaller than $f(t)$, as it is (for example) for negative powers of t. In this case we shall be able to say not only that the integral in the remainder term is small, but also how small it is. We may as well notice right now that the choice $a = 1/2$ is likely to be particularly convenient since the integral will then be the whole remainder term. However, if we do take $a = 1/2$ the limits on the integrals look awkward, and it is more usual to take $a = 0$. If we take $a = 0$ we get

$$\sum_{k=1}^{n} f(k) - \int_1^{n+1} f(t)dt = \tfrac{1}{2}\{f(1) - f(n+1)\} + \int_1^{n+1} P_1(t)f'(t)dt.$$

It is also usual to add $f(n+1)$ to both sides, so that we have

$$\sum_{k=1}^{n+1} f(k) - \int_1^{n+1} f(t)dt = \tfrac{1}{2}\{f(1) + f(n+1)\} + \int_1^{n+1} P_1(t)f'(t)dt. \tag{1.5}$$

We could now make the formula slightly tidier by replacing $n+1$ by n everywhere in (1.5).

It is sometimes useful to notice that even if f' does not exist we can still write (1.5) but with the integral on the right replaced by $\int P_1(t)df(t)$.

The case $a = 1/2$ produces

$$\sum_{k=1}^{n} f(k) - \int_{1/2}^{n+1/2} f(t)dt = \int_{1/2}^{n+1/2} P_1(t)f'(t)dt. \tag{1.6}$$

Formulas (1.5) and (1.6) are the simplest versions of the Euler–Maclaurin formula; usually the name refers to (1.5) (and its extensions that will be discussed later); (1.6) is called the second form of the Euler–Maclaurin formula. Let me remind you that we can replace the 1 in the lower limits and in $f(1)$ by any integer (less than n); in doing this in (1.6) we must think of the lower limit 1/2 as $1-1/2$. We can replace n by ∞ in (1.6) provided that two out of the three terms in (1.6) approach limits as $n \to \infty$.

Also we can replace the "step" 1 by s by applying (1.5) or (1.6) to $f(sx)$ instead of to $f(x)$. When $s = b/n$ this gives us

$$\sum_{k=0}^{n} f(kb/n) - (b/n)\int_0^b f(u)du = \tfrac{1}{2}\{f(b) + f(0)\} + (b/n)\int_0^b P_1(nu/b)f'(u)du.$$

If we solve this for the integral on the left we obtain the trapezoidal rule for approximating $\int_0^b f(u)du$ by a sum, together with the remainder term for the approximation.

Either of (1.4) or (1.5) contains the integral test for convergence of series; we have really just elaborated the usual proof of this test. The test is ordinarily stated as saying that the integral and the series converge or diverge together; we can say, in addition, how large the difference between them is. We could easily derive a more general version of the integral test: if f has bounded variation on some interval (m,∞) then $\int^\infty f$ and $\Sigma^\infty f(n)$ converge or diverge together [2]. This is true in particular if $\int^\infty |f'(t)|\,dt$ converges. For example, $\int^\infty t^{-1}\sin(t^{1/2})dt$ converges (by the argument usually given for the convergence of $\int^\infty t^{-1}\sin t\,dt$), the total variation of the integrand on $(1,\infty)$ is finite, and it follows that $\Sigma\, n^{-1}\sin(n^{1/2})$ converges, a fact which is not easy to prove in any completely different way.

When f decreases to 0, it is easy to see (again by the same argument as for $\int^\infty t^{-1}\sin t\,dt$) that

the left-hand side of (1.5) approaches a limit; when $f(x)=1/x$, this limit is Euler's constant, denoted by γ (sometimes by C). (Although γ has been computed to more than 7000 decimal places, beginning with 0.57721 56649, nothing much is known about it, not even whether or not it is rational.)

2. One can do a certain amount with the Euler–Maclaurin formulas in the versions that we now have (as we saw in Chapter I), but we can sometimes do much better by integrating by parts again to introduce higher derivatives of f. This is likely to be helpful provided that we can calculate several derivatives of f without too much trouble, that some derivative is small (in absolute value), and that $\int f(t)dt$ comes out in terms of tabulated functions. These are pretty restrictive requirements, but there are many interesting cases where they are all satisfied.

We are going to have to integrate P_1 repeatedly, and since P_1 has period 1 it seems like a good idea to use its periodic integrals. Since $P_1(x)$ is equal to a polynomial ($x-1/2$, in fact) on $(0,1)$, its successive integrals will be polynomials on $(0,1)$. We shall call these polynomials $Q_n(x)$, so that $Q_1(x)=P_1(x)$ on $(0,1)$, $Q_2'(x)=Q_1(x)$, and generally $Q_k'(x)=Q_{k-1}(x)$, $(k>1)$. By repeating $Q_k(x)$ with period 1 we get a periodic function $P_k(x)$. Since P_k is to be an integral of P_{k-1}, it had better be continuous when $k>1$. For continuity we must have $Q_k(1)=Q_k(0)$. To keep the formulas tidy, we define $Q_0(x)=1$; then $Q_1'(x)=Q_0(x)$, but $Q_1(1)\neq Q_1(0)$. Summing up, then, we have

$$\begin{aligned} &Q_k'(x)=Q_{k-1}(x), && k\geqq 1;\\ &Q_k(1)=Q_k(0), && k>1;\ Q_1(1)-Q_1(0)=1. \end{aligned}$$

(A set of polynomials satisfying the first equation is called an Appell set.)

Finally, for historical reasons, we define $B_k(x)=k!Q_k(x)$, and $B_k=B_k(0)=k!Q_k(0)$. The B's are called Bernoulli polynomials and Bernoulli numbers. We could compute as many as we need from the definition, but it is easier in the long run to get formulas for them. We shall do this in Chapter III.

Warning: notations vary. Our B_k's are the numbers so denoted in, for example, **[1]** and **[8]**. Several hundred of them have been computed **[15]**. The first few are

$$\begin{aligned} &B_0=1,\quad B_1=-\tfrac{1}{2},\quad B_2=\tfrac{1}{6},\quad B_4=-\tfrac{1}{30},\quad B_6=\tfrac{1}{42},\\ &B_8=-\tfrac{1}{30},\quad B_{10}=\tfrac{5}{66};\quad B_k=0 \text{ for odd } k>1. \end{aligned}\tag{2.1}$$

They do not look particularly complicated up to this point, but they get worse as they go on: for example, $B_{12}=-\frac{691}{2730}$, and $B_{20}=-\frac{174611}{330}$; the numerator of B_{60} has 43 decimal digits.

We can now integrate repeatedly by parts in (1.4). The integrated terms will be $P_k(1-a)$ and $P_k(n+1-a)$, which are equal since P_k has period 1. We have $P_k(1-a)=Q_k(1-a)=B_k(1-a)/k!$. The end result is

$$\begin{aligned} \int_{1-a}^{n+1-a} P_1(t)f'(t)dt &= \frac{B_2(1-a)}{2!}\{f'(n+1-a)-f'(1-a)\}-\frac{B_3(1-a)}{3!}\{f''(n+1-a)-f''(1-a)\}\\ &\quad+\cdots+\frac{B_{2m}(1-a)}{(2m)!}\{f^{(2m-1)}(n+1-a)-f^{(2m-1)}(1-a)\}\\ &\quad-\frac{B_{2m+1}(1-a)}{(2m+1)!}\{f^{(2m)}(n+1-a)-f^{(2m)}(1-a)\}\\ &\quad+\int_{1-a}^{n+1-a} f^{(2m+1)}(t)P_{2m+1}(t)dt. \end{aligned}\tag{2.2}$$

If we now insert (2.2) into (1.4) we have the general Euler–Maclaurin formula.

The usual versions come from $a=0$ and $a=1/2$. When $a=0$ all the terms involving B's with odd index drop out, and we get what is usually called *the* Euler–Maclaurin formula:

$$\sum_{k=1}^{n} f(k)-\int_1^n f(t)dt=\tfrac{1}{2}\{f(n)+f(1)\}+\frac{B_2}{2!}\{f'(n)-f'(1)\}$$
$$(2.3)\qquad +\cdots+\frac{B_{2m}}{(2m)!}\{f^{(2m-1)}(n)-f^{(2m-1)}(1)\}$$
$$+\int_1^n f^{(2m+1)}(t)P_{2m+1}(t)dt.$$

When $a=1/2$ the terms with odd index drop out again, since it can be shown that $B_k(1/2)=0$ for odd $k>1$; in fact, $B_k(1/2)=(2^{1-k}-1)B_k$. The resulting formula is the "second form" of the Euler–Maclaurin formula [**12**]:

$$\sum_{k=1}^{n} f(k)-\int_{1/2}^{n+1/2} f(t)dt=\frac{B_2(\frac{1}{2})}{2!}\{f'(n+\tfrac{1}{2})-f'(\tfrac{1}{2})\}$$
$$(2.4)\qquad +\cdots+\frac{B_{2m}(\frac{1}{2})}{(2m)!}\{f^{(2m-1)}(n+\tfrac{1}{2})-f^{(2m-1)}(\tfrac{1}{2})\}$$
$$+\int_{1/2}^{n+1/2} f^{(2m+1)}(t)P_{2m+1}(t)dt.$$

It is tempting to let $m\to\infty$ in (2.3) to obtain an infinite series for the left-hand side, but since the Bernoulli numbers increase quite fast it turns out that only very special kinds of function will have convergent series. It is possible to get badly misled if we replace the index 1 by $-n$ and then let $n\to\infty$; if f and all its derivatives tend to 0 at $\pm\infty$ then every term on the right-hand side is zero except for the integral. If we neglect the remainder (as we might do if we had used a formal, "operational," derivation of the formula), we would apparently have

$$\sum_{n=-\infty}^{\infty} f(n)=\int_{-\infty}^{\infty} f(t)dt.$$

This is not usually correct, although the two sides can be remarkably close ([**17a**], pp. 90–93). For example, when $f(x)=e^{-x^2}$, the five-decimal-place values of the integral and sum, respectively, are $\pi^{1/2}=1.77245$ and 1.77264 [**17**].

In Chapter I we saw how to use the simplest version of the Euler–Maclaurin formula to determine how fast a convergent series converges or how fast a divergent series diverges; in Chapter IV we shall see some applications of the more elaborate versions (2.3) or (2.4).

Right now I want to give a much simpler application. Remember that we said that the Euler–Maclaurin formula is especially useful when some derivative of f is quite small. It would be best of all if some derivative of f were 0, since then the integral in the remainder would drop out. This will happen whenever f is a polynomial. The simplest case is $f(x)=x^p$ with p a positive integer. Then (2.3) becomes

$$(2.5)\qquad \sum_{k=1}^{n} k^p=\frac{n^{p+1}}{p+1}+\frac{1}{2}n^p+\frac{B_2}{2!}pn^{p-1}+\frac{B_4}{4!}p(p-1)(p-2)n^{p-3}+\cdots,$$

ending with nB_p when p is even and with n^2pB_{p-1} when n is odd, since when p is odd the last term before the (zero) remainder is $B_{p+1}(p!-p!)/(p+1)!=0$.

Let's try $p=1$ and $p=2$. We get the familiar formulas

$$\sum_{k=1}^{n} k=\tfrac{1}{2}n^2+\tfrac{1}{2}n=\tfrac{1}{2}n(n+1),$$

$$\sum_{k=1}^{n} k^2=\tfrac{1}{3}n^3+\tfrac{1}{2}n^2+\tfrac{1}{6}n=\tfrac{1}{6}n(n+1)(2n+1).$$

For larger p and especially for large n, (2.5) can be impressively efficient even in the era of high-speed computers, provided that one wants an exact value of the sum (for what reason I do not really know). For example, if we take $p=9$ and $n=1000$ we find

$$\begin{aligned}\sum_{k=1}^{1000} k^9 &= \frac{1000^{10}}{10}+\tfrac{1}{2}(1000)^9+\frac{B_2}{2}\cdot 9\cdot(1000)^8+\frac{B_4}{4!}\cdot 9\cdot 8\cdot 7\cdot(1000)^6\\ &\quad+\frac{B_6}{6!}\cdot 9\cdot 8\cdot 7\cdot 6\cdot 5\cdot(1000)^4+\frac{B_8}{8!}\cdot 9\cdot 8\cdot\cdots\cdot 3\cdot(1000)^2\\ &=10^{29}+\tfrac{1}{2}\times 10^{27}+\tfrac{3}{4}\times 10^{24}-7\times 10^{17}+\tfrac{1}{2}\times 10^{12}-\tfrac{3}{2}\times 10^5\\ &=10050\ 07499\ 99300\ 00049\ 99998\ 50000.\end{aligned}$$

For the corresponding calculation with 10th powers see [**9**], p. 63. I have to admit that I chose the number of terms to show the formula to the best advantage; it would be considerably harder if n were, say, 563; but even in that case we would have to calculate only 6 different powers of 563 instead of one power of each of 563 integers. In point of fact, powers of three-digit integers can be looked up in tables [**16**], [**10**], and after a little arithmetic we get

$$\sum_{k=1}^{563} k^9 = 32\ 28004\ 07928\ 53932\ 99368\ 25476.$$

3. Most applications of the Euler–Maclaurin formula are to series of positive terms, but the formula has something to say even for alternating series of the form $\Sigma(-1)^{k-1}f(k)$. The obvious approach is to approximate the sums of the positive and negative terms separately by the Euler–Maclaurin formula and then subtract, but if both sums are numerically large we might lose a good deal of accuracy this way. The subtle way to do the job is to do the subtraction first and the computation afterwards. We can get a compact result, known as Boole's formula, by using (2.3) for the terms $f(2k)$ and (2.4) for the terms $f(2k-1)$. It turns out that the integrals on the left-hand side cancel each other, and we are left, after some algebra, with

$$\begin{aligned}\sum_{k=1}^{2n-1}(-1)^{k-1}f(k) &= \tfrac{1}{2}\{f(2n-1)+f(1)\}+(2-2^{-\frac{1}{2}})B_2\{f'(2n-1)-f'(1)\}\\ &\quad+\cdots+\frac{2^{2m-1}}{(2m)!}(2-2^{1-2m})B_{2m}\{f^{(2m-1)}(2n-1)-f^{(2m-1)}(1)\}\\ &\quad+\int_1^n f^{(2m+1)}(2t-1)\{P_{2m+1}(t)-P_{2m+1}(t-\tfrac{1}{2})\}\,dt.\end{aligned}\tag{3.1}$$

There is a similar formula if we stop with a term of even index.

By a similar method we could deal with any sum $\Sigma c_k f(k)$ with periodic $\{c_k\}$, for example

$$4f(1)+2f(2)+4f(3)+\cdots+4f(2n+1),$$

which corresponds to Simpson's rule in the same way that the Euler–Maclaurin formula corresponds to the trapezoidal rule (see [**3**]).

It is interesting to apply (3.1) (with $m=0$) to estimate the remainder in an alternating series. The textbook estimate is that when the terms decrease in absolute value and alternate in sign, the remainder (which obviously has the sign of the first term neglected) has absolute value less than the absolute value of the first term neglected (or of the last term retained). In the common case when f' is monotonic, this estimate is actually too weak by a factor of about 2: the absolute value of the remainder is less than half the absolute value of the last term retained. Let us verify this in the case where we stop with a term of odd index. (The other case is exactly parallel.) If we take $m=0$ in (3.1),

replace the lower limit 1 by $2n-1$, and let the upper limit become infinite, we get

$$\sum_{k=2n-1}^{\infty}(-1)^{k-1}f(k)=\tfrac{1}{2}f(2n-1)+2\int_{n}^{\infty}f'(2t-1)\{P_1(t)-P_1(t-\tfrac{1}{2})\}dt,$$

whence

$$\sum_{k=2n}^{\infty}(-1)^{k-1}f(k)=-\tfrac{1}{2}f(2n-1)+2\int_{n}^{\infty}f'(2t-1)\{P_1(t)-P_1(t-\tfrac{1}{2})\}dt. \tag{3.2}$$

Let us suppose that $f(t)\geqq 0$, $f'(t)<0$, and $|f'(t)|$ decreases. Then since $P_1(t)-P_1(t-1/2)$ is a "square wave" with values $-1/2$ on $(0,1/2)$ and $+1/2$ on $(1/2,1)$, and has period 1, it is clear that the integral is positive and the remainder (the left-hand side of (3.2)) is less in absolute value than $\frac{1}{2}f(2n-1)$, which is indeed half the last term retained (cf. [7]).

III. Bernoulli numbers. I have not yet made good on my claim that the Euler-Maclaurin formula will explain why Bernoulli numbers appear in so many power series for elementary functions. To do this I need better formulas for the Bernoulli numbers than we have had so far.

One of the standard methods for studying a sequence of numbers is to use them as coefficients in a power series, study the power series, and hence find properties of its coefficients. Instead of using the Bernoulli numbers themselves as coefficients, it is better to use the Bernoulli polynomials; and the formulas will be simpler if we work with the $Q_k(x)=B_k(x)/k!$. Accordingly we write down the power series

$$\sum_{k=0}^{\infty}Q_k(x)t^k. \tag{1}$$

To begin with, this is a completely formal series — we do not even know whether it converges for any values of t except $t=0$.

Recall that $Q_k'(x)=Q_{k-1}(x)$, so with a shift in indices $Q_{k+1}'(x)=Q_k(x)$. Since

$$\int_0^1 Q_k(t)dt = Q_{k+1}(1)-Q_{k+1}(0)=0$$

when $k\geqq 1$, there has to be a zero z_k of Q_k on $(0,1)$ so that $Q_k(t)$, which is continuous, can change sign. Using this z_k, we have

$$Q_k(x)=\int_{z_k}^{x}Q_{k-1}(t)dt. \tag{2}$$

Now $|Q_1(x)|=|P_1(x)|\leqq 1/2$ on $(0,1)$, so (2) with $k=2$ shows that $|Q_2(x)|\leqq 1/2$ for $0\leqq x\leqq 1$. Proceeding by induction, we obtain $|Q_k(x)|\leqq 1/2$ for $k\geqq 1$, $0\leqq x\leqq 1$. This is far from a precise inequality, but it is enough to show that the power series (1) converges at least for $|t|\leqq 1$. Since $Q_k'(x)=Q_{k-1}(x)$, the same is true for the series obtained by differentiating (1) with respect to x, any number of times.

Now we can write

$$\frac{d}{dx}\sum_{k=0}^{\infty}Q_k(x)t^k=\sum_{k=0}^{\infty}t^kQ_k'(x)=\sum_{k=1}^{\infty}t^kQ_{k-1}(x)=t\sum_{k=0}^{\infty}t^kQ_k(x),$$

at least for $|t|<1$. Hence for each t the series (1) converges to a solution of the differential equation $y'-ty=0$ (the differentiation is with respect to x). The solutions of $y'-ty=0$ are the same as the solutions of $e^{-xt}(y'-ty)=0$, and the latter equation says that

$$\frac{d}{dx}(ye^{-xt})=0.$$

But this last equation implies that $ye^{-xt} = A(t)$, that is, $A(t) = e^{-xt}\sum_{k=0}^{\infty} Q_k(x)t^k$, i.e.,

$$A(t)e^{xt} = \sum_{k=0}^{\infty} Q_k(x)t^k, \tag{3}$$

at least for $|t|<1$. Moreover, if we put $t=0$ we get $A(0) = Q_0(x) = 1$. We now have enough information to extract the explicit form of $A(t)$ from (3). Remember that for $k=0$ and for $k>1$ we have $Q_k(0) = Q_k(1)$; but $Q_1(1) - Q_1(0) = 1$. For $x=0$, (3) says

$$A(t) = \sum_{k=0}^{\infty} Q_k(0)t^k. \tag{4}$$

For $x=1$, it says

$$A(t)e^t = \sum_{k=0}^{\infty} Q_k(1)t^k = \sum_{k=0}^{\infty} Q_k(0)t^k + t = A(t) + t.$$

Hence we have

$$A(t) = A(t)e^t - t, \tag{5}$$

whence

$$A(t) = \frac{t}{e^t - 1}. \tag{6}$$

This lets us write (3) in the form

$$\frac{t}{e^t-1}e^{xt} = \sum_{k=0}^{\infty} Q_k(x)t^k = \sum_{k=0}^{\infty} \frac{B_k(x)}{k!}t^k. \tag{7}$$

We can use (7) to calculate the $Q_k(x)$ by the usual differentiation formulas for the coefficients in a power series. There is, however, an easier way. Some observant person noticed that

$$A(t) + \tfrac{1}{2}t = \frac{t}{e^t-1} + \frac{t}{2} = \tfrac{1}{2}t\frac{e^t+1}{e^t-1} = \tfrac{1}{2}t\frac{e^{\frac{1}{2}t}+e^{-\frac{1}{2}t}}{e^{\frac{1}{2}t}-e^{-\frac{1}{2}t}}. \tag{8}$$

From the last expression it is clear that $A(t)+\frac{1}{2}t$ is an even function; consequently the series (1) contains no odd powers except for the first one. This means that we must have $Q_k(0) = 0$ for odd $k>1$.

Now go back to (5), which we can write as $e^tA(t) = A(t) + t$, that is,

$$e^t\sum_{k=0}^{\infty} Q_k(0)t^k = \sum_{k=0}^{\infty} Q_k(0)t^k + t.$$

If we write $e^t = \sum_{n=0}^{\infty} t^n/n!$, multiply the two power series on the left, and compare powers of t on the two sides, we get

$$Q_k(0) = \sum_{j=0}^{k} Q_j(0)/(k-j)!, \qquad k>1.$$

This can be rewritten as

$$k!\,Q_k(0) = \sum_{j=0}^{k} \binom{k}{j} j!\,Q_j(0), \qquad k>1.$$

But $j!Q_j(0) = B_j$, so we have

$$B_k = \sum_{j=0}^{k} \binom{k}{j} B_j, \qquad k>1. \tag{9}$$

We can now start from $B_1 = -1/2$ and calculate the numbers B_j recursively. An easy way to remember (9) is to write it symbolically as

$$(B+1)^k - B^k = 0, \qquad k > 1, \tag{10}$$

where the first term in (10) is to be expanded by the binomial theorem and the exponents then degraded to subscripts. For other formulas for the B's see [11].

We may now calculate the values given on p. 248 for the B_k's, namely

$$B_1 = -\tfrac{1}{2}, \quad B_2 = \tfrac{1}{6}, \quad B_4 = -\tfrac{1}{30}, \quad B_6 = \tfrac{1}{42}, \quad B_8 = -\tfrac{1}{30}, \quad B_{10} = \tfrac{5}{66},$$

and so on; the B_k for odd $k > 1$ are, as we have already noticed, all zero. Unfortunately the notation for Bernoulli numbers is not uniform in the literature, but one can usually see what it is by comparing the first few B's as one finds them with the list above.

In a similar way we can find the polynomials $B_k(x)$. By (7) we have

$$\frac{t}{e^t - 1} e^{xt} = \sum_{k=0}^{\infty} \frac{B_k(x)}{k!} t^k.$$

This shows that

$$\begin{aligned} B_k(x) &= \left(\frac{d}{dt}\right)^k \{A(t)e^{xt}\}_{t=0}, \qquad A(t) = \frac{t}{e^t - 1}, \\ &= \sum_{j=0}^{k} \binom{k}{j} A^{(j)}(0) x^{k-j} = \sum_{j=0}^{k} \binom{k}{j} B_j x^{k-j}. \end{aligned}$$

Symbolically, this can be written as $B_k(x) = (B + x)^k$, with the same interpretation as in (10).

We are now ready to compute some power series. We need one additional piece of information, namely Euler's formulas for the trigonometric functions:

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i}, \qquad \cos x = \frac{e^{ix} + e^{-ix}}{2}.$$

We had

$$\sum_{k=0}^{\infty} \frac{B_k}{k!} t^k + \tfrac{1}{2}t = A(t) + \tfrac{1}{2}t = \tfrac{1}{2}t + \frac{e^{\frac{1}{2}t} + e^{-\frac{1}{2}t}}{e^{\frac{1}{2}t} - e^{-\frac{1}{2}t}}$$

for $|t| < 1$. In particular, this is true when $t = 2ix$:

$$ix + \sum_{k=0}^{\infty} \frac{B_k}{k!}(2ix)^k = ix\frac{e^{ix} + e^{-ix}}{e^{ix} - e^{-ix}} = x\frac{\cos x}{\sin x} = x \cot x.$$

Now $B_0 = 1$ and $B_1 = -\frac{1}{2}$, and $B_k = 0$ for odd $k > 1$, so

$$ix + 1 - ix + \sum_{k=1}^{\infty} \frac{B_{2k}}{(2k)!}(2ix)^{2k} = x \cot x,$$

$$\cot x = \frac{1}{x} + \sum_{k=1}^{\infty} \frac{B_{2k}}{(2k)!} 2^{2k}(-1)^k x^{2k-1},$$

$$\cot x = \frac{1}{x} - \frac{B_2}{2!}2^2 x + \frac{B_4}{4!}2^4 x^3 - \cdots.$$

If we integrate both sides, we get $\log \sin x$ in terms of Bernoulli numbers.

It is somewhat harder to get $\tan x$. We start from

$$\frac{1}{x+1} = \frac{1}{x-1} - \frac{2}{x^2 - 1}.$$

Hence with $x = e^t$ we have

$$\frac{t}{e^t+1} = \frac{t}{e^t-1} - \frac{2t}{e^{2t}-1}. \tag{11}$$

Now we know that

$$\frac{t}{e^t-1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} t^k, \quad \text{so} \quad \frac{2t}{e^{2t}-1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} 2^k t^k,$$

and by combining these in accordance with (11) we get

$$\frac{t}{e^t+1} = \sum_{k=0}^{\infty} \frac{B_k}{k!}(1-2^k)t^k.$$

Consequently if we replace t by $2t$ we have

$$\frac{2t}{e^{2t}+1} = \sum_{k=0}^{\infty} \frac{B_k}{k!}(1-2^k)2^k t^k.$$

But

$$\frac{2t}{e^{2t}+1} - t = t\frac{1-e^{2t}}{e^{2t}+1} = t\frac{e^{-t}-e^t}{e^t+e^{-t}},$$

so we have

$$t\frac{e^{-t}-e^t}{e^t+e^{-t}} = \sum_{k=0}^{\infty} \frac{B_k}{k!}(1-2^k)2^{2k}t^k - t.$$

Now take $t = ix$ and remember that $B_1 = -\frac{1}{2}$:

$$ix\frac{e^{-ix}-e^{ix}}{e^{ix}+e^{-ix}} = \sum_{k=2}^{\infty} \frac{B_k}{k!}(1-2^k)2^k i^k x^k.$$

The terms of odd index on the right drop out, and we get

$$\tan x = \sum_{k=1}^{\infty} \frac{B_{2k}}{(2k)!}(2^{2k}-1)2^{2k}(-1)^k x^{2k-1}.$$

Similarly we can get the expansion of $\csc t$ from

$$\frac{x}{x^2-1} = \frac{1}{x-1} - \frac{1}{x^2-1}.$$

The power series of $\sec t$, however, is not obtainable in this fashion.

We can get a quite different application of the Bernoulli numbers by forming the Fourier series of the periodic function $P_1(t)$ and integrating it term by term.

This process eventually produces the formula (see, for example, **[14]**)

$$\sum_{j=1}^{\infty} j^{-2k} = (-1)^{k-1}\frac{(2\pi)^{2k}}{2\cdot(2k)!}B_{2k}. \tag{12}$$

For example, $\sum_{j=1}^{\infty} j^{-2} = \pi^2/6$.

There are many other methods for summing the left-hand side of (12) **[19]**; but there is no known method of expressing $\sum_{j=1}^{\infty} 1/j^{2k+1}$ in "closed form" in terms of familiar numbers.

IV. Calculating sums and Euler constants. We have seen that for many convergent series it is impossible, even with machine calculation, to get a good numerical approximation to the sum of the series by adding up the successive terms. However, by using more elaborate forms of the

Euler–Maclaurin formula than we needed in Chapter I, we can evaluate the sums of some series, even though they may converge very slowly, to very high accuracy, with only a modest amount of calculation.

There is a similar problem if we want to calculate the Euler constant for a divergent series. We can discuss both problems together if we define the Euler constant for a convergent series by

$$\gamma = \sum_{k=1}^{\infty} f(k) - \int_1^{\infty} f(t)dt. \tag{1}$$

Knowing γ, we can, provided that the integral in (1) can be calculated, immediately calculate the sum of the series.

We shall first look at what can be done with the simplest version of the Euler–Maclaurin formula,

$$\sum_{k=1}^{n} f(k) - \int_1^{n} f(t)dt = \tfrac{1}{2}\{f(1) + f(n)\} + \int_1^{n} P_1(t)f'(t)dt. \tag{2}$$

Letting $n \to \infty$, we get

$$\gamma = \tfrac{1}{2}f(1) + \int_1^{\infty} P_1(t)f'(t)dt, \tag{3}$$

but we usually cannot calculate the integral directly. However, we can also write (2) in the form

$$\sum_{k=1}^{n} f(k) - \int_1^{n} f(t)dt = \tfrac{1}{2}\{f(1) + f(n)\} + \int_1^{\infty} P_1(t)f'(t)dt - \int_n^{\infty} P_1(t)f'(t)dt,$$

so that by (3)

$$\gamma = \sum_{k=1}^{n} f(k) - \int_1^{n} f(t)dt - \tfrac{1}{2}f(n) + \int_n^{\infty} P_1(t)f'(t)dt. \tag{4}$$

Here n is at our disposal, so we can take some convenient n, calculate the sum in (4) by hand (or machine), evaluate $\int f(t)dt$, and estimate the error by estimating the integral on the right. A crude but useful estimate comes from noticing that $|P_1(t)| \leqq 1/2$, so that

$$\left| \int_n^{\infty} P_1(t)f'(t)dt \right| \leqq \tfrac{1}{2} \left| \int_n^{\infty} f'(t)dt \right| = \tfrac{1}{2}f(n).$$

In the case when $f'(t) \leqq 0$ and $|f'|$ decreases, it is easy to see (as a sketch suggests) that since $P_1(t)$ has period 1, is negative on $(0, 1/2)$, and positive on $(1/2, 1)$, and symmetric about $t = 1/2$, we have

$$\int_n^{\infty} P_1(t)f'(t)dt \geqq 0.$$

Hence we have

$$\sum_{k=1}^{n} f(k) - \int_1^{n} f(t)dt - \tfrac{1}{2}f(n) \leqq \gamma \leqq \sum_{k=1}^{n} f(k) - \int_1^{n} f(t)dt. \tag{5}$$

Remember that (5) holds whether our series converges or diverges, γ being defined by (1) for a convergent series.

How much we can get out of (5) depends on how many terms of $\Sigma f(k)$ we are willing and able to add up directly. To get some idea of what happens in a specific case, let us try to calculate $\Sigma_1^{\infty} n^{-3} = \zeta(3)$. If we take $n = 10$ we can add up the sum by hand or on a pocket calculator, get γ, and hence get the sum of the series with an error that is certainly less than $1/2 \times 10^{-3}$. If we take $n = 10000$ and add up the sum by machine we will get γ with an error of less than $1/2 \times 10^{-12}$. This is a good deal easier than adding the 10^6 terms that would be needed to get this much accuracy by direct addition of

the terms one at a time. (Parenthetically, I ought to point out that — as anyone experienced with a computer is aware — one should always add up a long sum by starting at the end where the terms are smallest, in order not to lose too much accuracy from rounding errors.)

However, we can do much better by using one of the more refined versions of the Euler–Maclaurin formula. We can replace (4) by

$$(6)\quad \gamma = \sum_{k=1}^{n} f(k) - \int_1^n f(t)dt - \tfrac{1}{2}f(n) - \frac{B_2}{2!}f'(n) - \cdots - \frac{B_{2m}}{(2m)!}f^{(2m-1)}(n) + \int_n^\infty f^{(2m+1)}(t)P_{2m+1}(t)dt.$$

For series of the kind we have been considering, $f^{(2m+1)}(t)$ will, for an m of moderate size, tend to 0 rather fast as $t \to \infty$, so that by using a few Bernoulli numbers and adding up some terms at the beginning of the series, we can get a quite accurate approximation to γ. As an example, let us take $f(x) = x^{-3}$ again. In our previous calculation we got an error of less than 10^{-3} by adding 10 terms of the series and using (6) with $m = 0$. If we take $m = 3$, the error can be shown (see the appendix to this chapter) to be at most 1.5×10^{-11}, which means that with 10 terms of the series and 3 Bernoulli numbers, we can get almost as much accuracy as we got before with 10^4 terms. If we do use 10^4 terms we now get an upper bound of 1.5×10^{-41} for the error.

We can treat divergent series similarly. For example, for the harmonic series 10 terms and 3 Bernoulli numbers give γ with an error of less than $1/2 \times 10^{-10}$.

Once we know γ, we can get compact formulas for the partial sums of either convergent or divergent series. For example, we have

$$\sum_{k=1}^{N} f(k) = \gamma + \tfrac{1}{2}f(N) + \int_1^N f(t)dt + \tfrac{1}{12}f'(N) - \tfrac{1}{720}f'''(N) - \int_N^\infty P_5(t)f^{(5)}(t)dt.$$

For the harmonic series this becomes

$$\sum_{k=1}^{N} \frac{1}{k} = \log N + \gamma + \frac{1}{2N} - \frac{1}{12N^2} + \frac{1}{120N^4} + 120\int_N^\infty P_5(t)t^{-6}dt,$$

which represents the partial sums with an error that decreases rapidly as N increases — in fact, the error can be shown to be at worst $0.004N^{-6}$. Consequently we can compute partial sums with considerable accuracy provided that we have good approximations to γ and $\log N$.

Appendix to Chapter IV. In many cases we can get quite accurate estimates of remainders in the following way ([**14**], pp. 536ff.). Most of the series that one wants to work with have terms $f(n)$ coming from a function whose successive derivatives $f^{(n)}(x)$ eventually alternate in sign, at least for large values of n and large x, and tend to 0 as $x \to \infty$. What makes these functions particularly relevant for our purposes is that practically any function concocted from powers and logarithms, and tending to 0 at ∞, will have this property ([**13**], p. 17).

Now recall that the remainder in the Euler–Maclaurin formula (6) after the term in $f^{(2m-1)}(n)$ is

$$R_m = \int_n^\infty f^{(2m+1)}(t)P_{2m+1}(t)dt.$$

We first have to know something about what the periodic functions P_{2m+1} look like over a period. We are already familiar with P_1 (Figure 3), for which $P_1(x) = x - 1/2$ on $(0,1)$. Now P_2 (Figure 4) is the integral of P_1 that has average value 0 and is equal to B_2 at $x = 0$ and 1, so $P_2(x) = 1/2x^2 - 1/2x + \frac{1}{12}$, and has two zeros on $(0,1)$. Moreover, since P_1 is odd with respect to the point $x = 1/2$, the function P_2 is even with respect to the same point. Hence P_3 is again odd with respect to $x = 1/2$, and so is zero there; and since $P_3(1) = P_3(0)$ by periodicity, and $P_3(0) = B_3 = 0$ because the Bernoulli numbers of odd index > 1 are all zero, we have $P_3 = 0$ at 0 and 1. Similarly P_4 is even with respect to $x = 1/2$.

Now P_3 has at least 3 zeros in a period (at $0, 1/2, 1$); if it had more than 3 zeros, its derivative P_2 would have more than 2, which it doesn't. Hence P_3 has precisely the 3 zeros we have located.

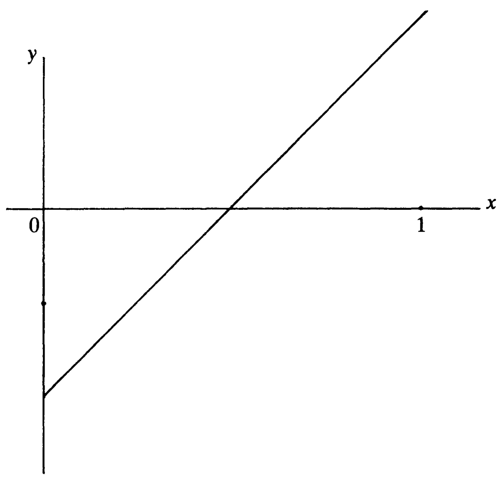


FIG. 3

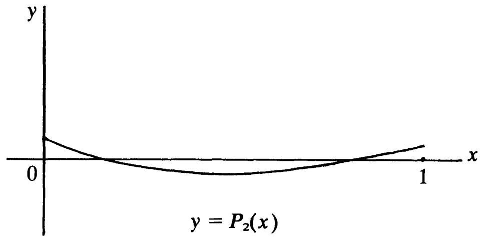


FIG. 4

Similarly P_2 has just the two zeros that the graph indicates. Now consider P_4. Its derivative P_3 has only one zero inside $(-1,1)$, so P_4 has at most two zeros between -1 and 1. Moreover, if $P_4(0)$ or $P_4(1)$ were 0, P_3 would have another zero, which it doesn't (and of course not, since $P_4(0)=B_4\neq 0$). Continuing in this way, we can see that P's of odd index greater than 2 have zeros at 0, 1/2 and 1, and are odd around $x=1/2$; P's of even index are even around $x=1/2$ and have two interior zeros. We can also see that P_{4k+1} is negative on $(0,1/2)$ whereas P_{4k+3} is positive there; and P_{4k+2} is positive at 0 and 1, whereas P_{4k} is negative at these points.

Now suppose again that $|f^{(2m+1)}|$ decreases. Since $P_{2m+1}(t)$ has period 1, and is symmetric on each $(n,n+1)$ about the point $n+1/2$, it follows that if (for example) $f^{(2m+1)}(t)$ is positive and decreasing and $P_{2m+1}(t)>0$ on $(0,1/2)$, and hence on each $(n,n+1/2)$, we have $R_m=\int_n^\infty f^{(2m+1)}(t)P_{2m+1}(t)dt>0$; in this case we are assuming $f^{(2m+3)}(t)>0$, but $P_{2m+3}(t)<0$ on $(n,n+1/2)$, and therefore $R_{m+1}<0$. Consequently $R_m-R_{m+1}>R_m$. A similar inequality holds for each of the other possible choices of signs for $f^{(2m+1)}$ and P_{2m+1}, and we find that in all cases R_m-R_{m+1} has the sign of R_m, and also

$$|R_m|<|R_m-R_{m+1}|.$$

Now look at formula (6), which expresses γ as a sum of terms ending with

$$-\frac{B_{2m}}{(2m)!}f^{(2m-1)}(n),$$

plus a remainder R_m. Taking the next larger m, we have γ expressed as the same sum followed by

$$-\frac{B_{2m+2}}{(2m+2)!}f^{(2m+1)}(n)+R_{m+1}.$$

Subtracting, we get

$$R_m-R_{m+1}=\frac{B_{2m+2}}{(2m+2)!}f^{(2m+1)}(n), \tag{7}$$

so that R_m has the same sign as, and a smaller absolute value than, the right-hand side of (7).

By way of illustration, I verify some numerical values that were quoted earlier in this chapter.

(a) Harmonic series, N terms ($n=N$), $m=3$ (3 Bernoulli numbers). The error in computing γ by

taking this many terms is between 0 and

$$-\frac{B_8}{8!}\frac{(-7!)}{N^8}=\frac{1}{240}N^{-8}.$$

For $N = 10$ this is less than $1/2 \times 10^{-10}$.

(b) $\Sigma\, n^{-3}$, $m = 3$, $n = N$. The error is at most $\frac{3}{20}M^{-10}$; so for $N = 10$ it is at most 0.15×10^{-10} and for $N = 10^4$ it is at most 0.15×10^{-40}.

References

1. M. Abramowitz and I. Stegun, editors, Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables, National Bureau of Standards, Applied Mathematics Series, 55, U.S. Government Printing Office, Washington, D.C., 1964.

2. R. P. Agnew and R. P. Boas, Jr., An integral test for convergence, this MONTHLY, 56 (1949) 677–678.

3. B. C. Berndt and L. Schoenfeld, Periodic analogues of the Euler–Maclaurin and Poisson summation formulas with applications to number theory, Acta Arith., 28 (1975) 23–68.

3a. L. Berg, Asymptotische Darstellungen und Entwicklungen, VEB Deutscher Verlag der Wissenschaften, Berlin, 1968.

4. R. P. Boas, Jr., Growth of partial sums of divergent series, Math. of Comp., 31 (1971).

5. ——— and J. W. Wrench, Jr., Partial sums of the harmonic series, this MONTHLY, 78 (1971) 864–870.

6. ——— and ———, unpublished calculations.

7. P. Calabrese, A note on alternating series, this MONTHLY, 69 (1962) 215–217.

8. L. Comtet, Advanced Combinatorics, revised ed., Reidel, Dordrecht and Boston, 1974.

8a. ———, Inversion de $y^{\alpha}e^{y}$ et $y \log^{\alpha} y$ au moyen des nombres de Stirling, C.R. Acad. Sci. Paris, 270 (A) (1970) 1085–1088.

9. H. T. Davis, The Summation of Series, Principia Press of Trinity University, San Antonio, Tex., 1962.

10. J. W. L. Glaisher, W. G. Bickley, C. E. Gwyther, J. C. P. Miller and E. J. Ternouth, Table of Powers Giving Integral Powers of Integers (British Association for the Advancement of Science, Mathematical Tables, vol. IX), Cambridge University Press, 1940.

11. H. W. Gould, Explicit formulas for Bernoulli numbers, this MONTHLY, 79 (1972) 44–51.

12. ——— and W. Squire, Maclaurin's second formula and its generalization, this MONTHLY, 70 (1963) 44–52.

13. G. H. Hardy, Orders of Infinity. The 'Infinitärcalcul' of Paul du Bois-Reymond, 2nd ed. Cambridge University Press, London, 1924.

14. K. Knopp, Theory and Application of Infinite Series, Blackie, London and Glasgow, 1928.

15. D. E. Knuth and T. J. Buckholtz, Computation of tangent, Euler, and Bernoulli numbers, Math. of Comp., 21 (1967) 663–688.

16. A. N. Lowan et al., Tables of the First Ten Powers of the Integers from 1 to 1000, Project for the Computation of Mathematical Tables, New York, 1939.

17. J. van de Lune, Enkele opmerkingen over het gebruik van de sommatieformule van Euler (Some remarks on the use of the Euler summation formula), Math. Centrum Amsterdam, Afd. Zuivere Wisk., 1965, ZW–011.

17a. J. N. Lyness, Computational techniques based on the Lanczos representation, Math. of Comp., 28 (1974) 81–123.

18. J. Miller and R. P. Hurst, Simplified calculation of the exponential integral, Math. Tables Aids Comp., 12 (1958) 187–193.

19. E. L. Stark, The series $\Sigma_{k=1}^{\infty} k^{-s}$, $s = 2, 3, 4, \ldots$, once more, Math. Mag., 47 (1974) 197–202.

20. A. D. Wadhwa, An interesting subseries of the harmonic series, this MONTHLY, 82 (1975) 931–933.

21. D. V. Widder, The Laplace Transform, Princeton University Press, 1941.

22. ———, Advanced Calculus, 2nd ed., Prentice-Hall, Englewood Cliffs, N.J., 1961.

DEPARTMENT OF MATHEMATICS, NORTHWESTERN UNIVERSITY, EVANSTON, IL 60201.

CAN ONE SEE THE SHAPE OF A SURFACE?

REESE T. PROSSER

Some ten years ago Professor Mark Kac published in these pages an article **[3]** entitled "Can one hear the shape of a Drum?". Under this provocative title Kac raised a fascinating question: To what extent is the shape of a smooth closed surface (a drum) determined by the eigenvalues of the classical Dirichlet problem for the interior? Since these eigenvalues are just the resonance frequencies for the acoustic wave motion in the interior, the question asks: to what extent is the shape of a surface determined by the sound of its interior?

This problem is a representative example of a class of problems, called *inverse* problems, which are quite different in character from the more familiar *direct* problems of partial differential equations. Generally speaking, direct problems ask us to develop a solution from the data, whereas inverse problems ask us to recover the data from the solution. Inverse problems have attracted a considerable and increasing interest in recent years because of their many immediate practical applications in our physical world. A splendid introduction may be found in Prof. Keller's recent article **[20]**. Typically they are very hard problems; Kac's problem, for example, remains unsolved today, though it has given rise to a considerable literature.

Here we should like to raise a companion question: To what extent is the shape of a surface determined by the eigenfunctions of the classical Dirichlet problem for the *exterior*?

We must note first of all that information about the shape of the drum is not to be found in the *eigenvalues* of the exterior problem, since, as we shall see, these consist of all negative real numbers, no matter what the shape. So whatever information about the shape can be recovered from the solution must be somehow contained in the *eigenfunctions* of the exterior problem. In particular, we expect on the basis of practical experience that this information should be available in the *asymptotic* behavior of these eigenfunctions at great distances from the surface, since it is this behavior that determines what we "see" when the surface is "illuminated" by a given "source" of light, radar, or sonar waves. That is, we expect to find that the shape of a surface is determined by the sight of its exterior. Thus we can see the shape of a drum, even if we can't hear it.

The direct problem. Let S be a smooth closed surface in $\mathbb{R}^3$. We shall always suppose that S is connected and compact. S then separates $\mathbb{R}^3$ into an interior region B and an exterior region B'. We shall search for solutions of the *exterior* Dirichlet problem:

$$\nabla^2 u(x) + \lambda^2 u(x) = 0 \qquad x \in B' \tag{1}$$

$$u(x) = 0 \qquad x \in S \tag{2}$$

in the form

$$u(x) = u_1(x) + u_2(x), \tag{3}$$

where $u_1(x)$ is a predetermined solution of (1) in *all* of $\mathbb{R}^3$ and $u_2(x)$ satisfies the *radiation condition* **[11]**

$$\left| r\left(\frac{\partial}{\partial r} - i\lambda\right) u_2(x) \right| \to 0 \quad \text{as} \quad r = |x| \to \infty. \tag{4}$$

We can think of $u_1(x)$ as the *illuminating* wave and $u_2(x)$ as the *scattered* wave for the problem. Note that if $u_1(x) \equiv 0$, then necessarily $u_2(x) \equiv 0$, so that we see nothing in the dark!

A representation of the solution. Since $u_1(x)$ is presumed known and at our disposal, it suffices to find $u_2(x)$. As a first step we can always represent $u_2(x)$ as an integral over S. This follows almost immediately from the old familiar identity of Green, valid for any region A with smooth compact

boundary ∂A:

$$(5)\qquad \int_A (u\nabla^2 v - v\nabla^2 u)dy = \int_{\partial A} \left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n}\right)dy.$$

If we choose for A the region lying between S and a large circumscribed sphere of radius R, and choose for u the scattered wave $u_2(\boldsymbol{y})$ and for v the Green's function

$$(6)\qquad G(\boldsymbol{x},\boldsymbol{y}) = \frac{e^{i\lambda|\boldsymbol{x}-\boldsymbol{y}|}}{4\pi|\boldsymbol{x}-\boldsymbol{y}|},$$

then we find

$$(7)\qquad u_2(\boldsymbol{x}) = \int_S \left(u_2\frac{\partial G}{\partial n} - G\frac{\partial u_2}{\partial n}\right)dy - \int_{|y|=R} \left(u_2\frac{\partial G}{\partial n} - G\frac{\partial u_2}{\partial n}\right)dy.$$

Now Wilcox [16] has shown that, because of (4), both u_2 and G satisfy decay conditions as $r = |\boldsymbol{x}| \to \infty$:

$$(8)\qquad \begin{aligned} |ru_2(\boldsymbol{x})| &\to 0 \\ |r^2\nabla u_2(\boldsymbol{x})| &\to 0. \end{aligned}$$

It follows that as $R \to \infty$ the second term on the right in (7) drops out, and we are left with the representation

$$(9)\qquad \begin{aligned} u_2(\boldsymbol{x}) &= \int_S \left(u_2\frac{\partial G}{\partial n} - G\frac{\partial u_2}{\partial n}\right)dy \qquad \boldsymbol{x} \in B' \\ 0 &= \int_S \left(u_2\frac{\partial G}{\partial n} - G\frac{\partial u_2}{\partial n}\right)dy \qquad \boldsymbol{x} \in B. \end{aligned}$$

This representation gives us u_2 in B' in terms of its values and those of its normal derivative on the boundary S, in strict analogy with the familiar special case of exterior potential theory ($\lambda = 0$).

This representation for $u_2(x)$, however, is not unique, a well-kept professional secret leading to considerable confusion for amateurs in the subject! To see this best, let's introduce a new function $u_3(\boldsymbol{x})$ satisfying

$$(10)\qquad \begin{aligned} (\nabla^2 + \lambda^2)u_3(\boldsymbol{x}) &= 0, \qquad \boldsymbol{x} \in B \\ u_3(x) &= 0, \qquad \boldsymbol{x} \in B'. \end{aligned}$$

The same argument which gives (9), when applied to B, will give

$$(11)\qquad \begin{aligned} u_3(\boldsymbol{x}) &= -\int_S \left(u_3\frac{\partial G}{\partial n} - G\frac{\partial u_3}{\partial n}\right)dy, \qquad \boldsymbol{x} \in B \\ 0 &= \int_S \left(u_3\frac{\partial G}{\partial n} - G\frac{\partial u_3}{\partial n}\right)dy, \qquad \boldsymbol{x} \in B'. \end{aligned}$$

Adding (9) and (11), we find

$$(12)\qquad u_2(\boldsymbol{x}) + u_3(\boldsymbol{x}) = \int_S \left(\mu(\boldsymbol{y})\frac{\partial G}{\partial n} - G\nu(\boldsymbol{y})\right)dt, \qquad \boldsymbol{x} \in B \cup B',$$

where now $\mu(\boldsymbol{y}) = u_2(\boldsymbol{y}) - u_3(\boldsymbol{y})$ and $\nu(\boldsymbol{y}) = (\partial u_2/\partial n)(\boldsymbol{y}) - (\partial u_3/\partial n)(\boldsymbol{y})$. The freedom of choice of $u_3(\boldsymbol{x})$ permits different versions of (9). For example, if u_3 is chosen so that $u_3 = u_2$ on S, then $\mu(\boldsymbol{y}) = 0$, and

$$u_2(x) = u_2(x) + u_3(x) = -\int_S G\nu dy, \quad x \in B', \tag{13}$$

while if u_3 is chosen so that $\partial u_3/\partial n = \partial u_2/\partial n$ on S, then $\nu(y) = 0$ and

$$u_2(x) = \int_S \mu \frac{\partial G}{\partial n} dy, \qquad x \in B'. \tag{14}$$

For the Dirichlet problem before us, it is traditional to choose the form (14) and solve for the difference $\mu = u_2 - u_3$ on S by the standard Fredholm method of potential theory: as $x \to z \in S$ in (14) we find [cf. **21**]

$$u_2(z) = \tfrac{1}{2}\mu(z) + \int_S \frac{\partial G}{\partial n}\mu dy = -u_1(z) = \text{known}. \tag{15}$$

This gives us an integral equation for $\mu(y)$ which can generally be solved for $\mu(y)$ by the Fredholm theory. But we may also choose the form (13) and solve for the difference $\nu = (\partial u_2/\partial n) - (\partial u_3/\partial n)$ on S, since as $x \to z \in S$

$$\frac{\partial u_2}{\partial n}(z) = \tfrac{1}{2}\nu(z) - \int_S G\nu dy = -\frac{\partial u_1}{\partial n}(z) = \text{known}. \tag{16}$$

One or the other of these two approaches may fail for particular values of λ because of the Fredholm alternative, but not both at once; so that we are always assured of a solution u_2 for the scattered wave in the form (12) for any choice of λ (cf. Wilcox, [17]).

A spherical expansion of the solution. Since we are especially interested in the behavior of the scattered wave at large distances, our next step is to expand $u_2(x)$ in powers of $1/r$ and examine the leading terms. To this end we have only to expand the Green's function $G(x, y)$ in powers of $1/r$ and incorporate the result in (12). Any standard textbook in EM theory [e.g., **13**] will tell us that

$$G(x,y) = \frac{i\lambda}{4\pi} h_0(\lambda|x-y|) = \frac{i\lambda}{4\pi} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} C_{mn} Y_{mn}(\boldsymbol{\theta}) Y_{mn}(\boldsymbol{\phi}) h_n(\lambda|x|) j_n(\lambda|y|). \tag{17}$$

Here the C_{mn} are known coefficients, Y_{mn} are known spherical harmonics of the unit vectors $\boldsymbol{\theta} = x/|x|$ and $\boldsymbol{\phi} = y/|y|$, and h_n and j_n are spherical Bessel functions whose asymptotic behavior was all worked out long ago. Since this expansion converges uniformly in all the variables present provided that $|x| \geqq |y| + \varepsilon$, we can insert it immediately in (12) and get:

$$u_2(x) = \frac{i\lambda}{4\pi} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} F_{mn}(\lambda) Y_{mn}(\boldsymbol{\theta}) h_n(\lambda|x|), \tag{18}$$

with coefficients $F_{mn}(\lambda)$ given by

$$F_{mn}(\lambda) = C_{mn} \int_S Y_{mn}(\boldsymbol{\phi}) \left(j_n(\lambda|y|)\nu(y) - \frac{\partial}{\partial n} j_n(\lambda|y|)\mu(y) \right) dy. \tag{19}$$

This expansion converges uniformly in any region where $|x| \geqq \sup_{y \in S}|y| + \varepsilon$, i.e., outside of any open sphere containing S.

The $F_{mn}(\lambda)$ are sometimes called the *spherical moments* of the scattered wave u_2. Their interest for us stems from the fact that they clearly must contain whatever information about the shape of S is available in u_2.

A far field expansion of the solution. Next we recall that the spherical Bessel functions $h_n(z)$ all

have the form [cf. **13**]:

$$h_n(z)=\frac{e^{iz}}{iz}p_n\left(\frac{1}{iz}\right), \tag{20}$$

where p_n is a polynomial of degree n. Hence we can rearrange (18) to give a series in $1/r$:

$$u_2(\boldsymbol{x})=\frac{e^{i\lambda r}}{4\pi r}\sum_{k=0}^{\infty} F_k(\boldsymbol{\theta},\lambda)(\lambda r)^{-k}, \tag{21}$$

where

$$F_k(\boldsymbol{\theta},\lambda)=\sum_{n=0}^{\infty}\sum_{m=-n}^{n} D_{kmn}F_{mn}(\lambda)Y_{mn}(\boldsymbol{\theta}), \tag{22}$$

with known coefficients D_{kmn}. In particular, the leading term ($k=0$) in (22) is determined by

$$F_0(\boldsymbol{\theta},\lambda)=\sum_n\sum_m i^{-n}C_{mn}F_{mn}(\lambda)Y_{mn}(\boldsymbol{\theta}). \tag{23}$$

The expression $F_0(\boldsymbol{\theta},\lambda)$ is sometimes called the *radiation pattern* of the scattered wave; it determines what is observed experimentally in u_2 at large distances from S. Equation (23) gives an expansion for the radiation pattern in terms of spherical harmonics, in which the coefficients are essentially the spherical moments.

Properties of the spherical moments. It follows from their definition in (19) that the spherical moments are well-behaved functions of λ. In fact, from known properties of the spherical Bessel functions involved, we infer that the $F_{mn}(\lambda)$ are analytic in λ in the whole range of $0<\lambda<\infty$, with the following asymptotic behavior [cf. **13**]:

$$F_{mn}(\lambda)\to\lambda^n f_{mn}(0)+O(\lambda^{n+1}) \qquad \lambda\to 0, \tag{24}$$

$$F_{mn}(\lambda)\to\lambda^{-1}f_{mn}(\infty)+O(\lambda^{-2}) \qquad \lambda\to\infty. \tag{25}$$

Here $f_{mn}(0)$ are the spherical moments of the corresponding potential problem ($\lambda=0$), and the $f_{mn}(\infty)$ are simply the leading terms at high frequency:

$$f_{mn}(0)=C_{mn}\int_S Y_{mn}(\boldsymbol{\phi})\left(\frac{1}{|\boldsymbol{y}|^n}\nu|\boldsymbol{y}|-\frac{\partial}{\partial n}\frac{1}{|\boldsymbol{y}|^n}\mu(\boldsymbol{y})\right)d\boldsymbol{y}, \tag{26}$$

$$\begin{aligned} f_{mn}(\infty)=C_{mn}\int_S Y_{mn}(\boldsymbol{\phi})&\left(\frac{\sin(\lambda|\boldsymbol{y}|-(n\pi/2))}{|\boldsymbol{y}|}\right)\nu(\boldsymbol{y})\\ &-\frac{\partial}{\partial n}\left(\frac{\sin(\lambda|\boldsymbol{y}|-n\pi/2)}{|\boldsymbol{y}|}\mu(\boldsymbol{y})\right)d\boldsymbol{y}. \end{aligned} \tag{27}$$

Note also that these moments are independent of the choice of μ and ν in the representation (19), since they depend only on the scattered field u_2 as shown by (22) and (23). Finally, if the illuminating field $u_1(x,\lambda)$ is a function only of the combination λx, as is often the case in practice, then it follows from (19) that both the illuminating and the scattered waves, and hence the moments, are invariant under the simultaneous scale change $\boldsymbol{x}\to b\boldsymbol{x}$ and $\lambda\to b^{-1}\lambda$. It follows that if a is the *diameter* of the surface S, then the spherical moments depend only on the combination λa.

The inverse problem. Suppose now that we are given an illuminating wave u_1 and a scattered wave u_2 and we seek to determine the shape of S. Without loss of essentials we may assume that the illuminating wave u_1 has the form of a plane wave traveling down the z-axis (i.e., from an overhead light).

$$u_1(\boldsymbol{x})=\exp i\lambda\boldsymbol{\omega}\cdot\boldsymbol{x}, \qquad \boldsymbol{\omega}=(0,0,-1). \tag{28}$$

On the basis of our knowledge of the direct problem, we shall deal with the inverse problem through a series of related steps.

Does the scattered field determine the surface? Fix λ and consider the function $u = u_1 + u_2$. We know that both u_1 and u_2 are analytic functions of r outside S (u_2 because of its representation) and that the sum $u = u_1 + u_2$ must vanish on S. Can u vanish on any other closed surface S'? Well, if it does, then S' cannot lie too far away from S, because outside of large spheres about S, $|u_1| = 1$ and $|u_2| \ll 1$. So S' must be bounded, and hence compact. Moreover, the interior of S' cannot be disjoint from the interior of S, since otherwise we can solve the direct problem for S' and get another function u' analytic outside S' and agreeing with u in the common exterior region. Hence u and u' together define a single function analytic throughout $\mathbb{R}^3$ and satisfying (4), and we know such a function must vanish identically. If at least some part of S' lies outside of S, then in the region C lying outside of S and inside of S' the function u satisfies the equation (1) and vanishes on the boundary. Unfortunately there is nothing in the theory of the direct problem to rule out this possibility: such a solution is simply an eigenfunction for C with eigenvalue λ. We know that there can be at most countably many eigenvalues for C with finitely many eigenfunctions for each. Hence, if we change the direction $\boldsymbol{\omega}$ of the illuminating wave u_1 (i.e., move the lamp) or change the value of λ (i.e., change the color) through any continuous range of values, we must find a solution that does not vanish on S'.

We conclude that the scattered field u_2 does determine the shape of the surface as the zero set of the function $u = u_1 + u_2$, up to the presence of "ghosts" which can always be eliminated by changing the direction or the frequency of the illumination. Under such a change the object S will remain fixed but the ghost S' will appear to move, [cf. **15**].

Do the spherical moments determine the scattered field? Yes, they do, via the spherical expansion (18) outside of the smallest sphere containing S. Inside this sphere u_2 is determined from its values outside by analytic continuation.

Does the radiation pattern determine the spherical moments? Yes, of course, via the expansion (23), since the coefficients in any orthogonal expansion, such as (23), can always be determined, and in fact determined by quadratures, from a knowledge of the sum.

Do experimental measurements determine the radiation pattern? Yes, they do, at least in principle. The *magnitude* of the radiation pattern has a simple interpretation: its square $\sigma(\boldsymbol{\theta}, \lambda) = |F_0(\boldsymbol{\theta}, \lambda)|^2$, sometimes called the *differential cross section*, is proportional to the intensity of the radiation scattered in the direction $\boldsymbol{\theta}$, and is determined by intensity measurements. The *argument* $\delta(\boldsymbol{\theta}, \lambda) = \arg F_0(\boldsymbol{\theta}, \lambda)$ of the radiation pattern is interpreted as a shift in phase of the original illumination due to the scattering process, and is determined by interference measurements. Finally, $F_0(\boldsymbol{\theta}, \lambda) = \sigma^{1/2}(\boldsymbol{\theta}, \lambda) e^{i\delta(\boldsymbol{\theta}, \lambda)}$.

Practical considerations. Thus we can say that in principle we can "see" the shape of a drum. In practice, however, our solution is not very useful as a constructive algorithm, since it involves the process of analytic continuation of the scattered wave from a sphere containing the surface back to the surface itself. This process will be highly unstable against small errors in the measurements of the radiation pattern unless S itself is nearly spherical. It is highly desirable, therefore, to consider whether there is a constructive algorithm for implementing our solution.

Low frequency approximation. If the frequency is so low that $\lambda a \ll 1$, then the scattered field is a slight perturbation of the static potential field. In particular, the spherical moments $F_{mn}(\lambda)$ of u_2 may be used to determine the static moments $f_{mn}(0)$ via (24). But if we know the static moments, then we can reconstruct the static potential for S from its moment expansion, and we can recover S from the static potential as the set where the potential achieves its *maximal* value of 1, [**21**]. Moreover, level sets of the potential below 1 will give an approximation for the shape of S which will give an idea of its

general outline, and some of these level sets may be entirely outside the smallest sphere containing S and hence may be directly computable from the moment expansion.

We note that the first spherical moment $f_\infty(0)$ is just equal to the (static) *capacity* of S, so that this capacity can be inferred directly from the radiation pattern:

$$f_\infty(0) = \lim_{\lambda \to 0} \int_{|\boldsymbol{\theta}|=1} F(\boldsymbol{\theta}, \lambda)\, d\boldsymbol{\theta}. \tag{29}$$

High frequency expansions. At the other end of the scale, if the frequency is so high that $\lambda a \gg 1$, the scattered field is a perturbation of the geometric optics limit obtained by the familiar ray-tracing arguments of geometric optics [**4**]. In order to see how this comes about, we shall need to know how to evaluate integrals like those in (27) when λ is very large.

The method of stationary phase. Every such integral can be reduced after suitable partitions of unity and changes of variable (and careful attention to the resulting Jacobians!) to the form

$$I(\lambda) = \int_{-\infty}^{+\infty} e^{i\lambda x^2/2} \phi(x)\, dx, \tag{30}$$

where $\phi(x)$ is a smooth function of compact support. Now if λ is very large then we expect that the contribution to $I(\lambda)$ from any interval not containing the origin will be very small, since the rapid oscillations of the exponential will tend to average to zero. The contribution from intervals about the origin, however, need not be so small because there the oscillations slow down as the phase becomes stationary. To evaluate $I(\lambda)$ it is simplest to take Fourier Transforms:*

$$I(\lambda) = \int_{-\infty}^{+\infty} \sqrt{\frac{i}{\lambda}} \exp\left(-\frac{iy^2}{2\lambda}\right) \hat{\phi}(y)\, dy \tag{31}$$

and expand the exponential in a power series in $1/\lambda$. Thus we find

$$\begin{aligned} I(\lambda) &\sim \sqrt{\frac{i}{\lambda}} \sum_{k=0}^{\infty} \frac{1}{k!} \int_{-\infty}^{+\infty} \left(-\frac{iy^2}{2\lambda}\right)^k \hat{\phi}(y)\, dy \\ &\sim \sqrt{\frac{i}{\lambda}} \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{i}{2\lambda}\right)^k \phi^{(2k)}(0). \end{aligned} \tag{32}$$

This series gives the asymptotic behavior of $I(\lambda)$ as $\lambda \to \infty$. In particular the leading term is determined solely by the value of ϕ at the origin:

$$I(\lambda) = \sqrt{\frac{i}{\lambda}} \phi(0) + O\left(\frac{1}{\lambda^{3/2}}\right). \tag{33}$$

We now turn to the evaluation of the $F_{mn}(\lambda)$ for large λ.

The physical optics approximation for convex surfaces. We begin by reconsidering the equation (15) which determines the value of $\mu(\boldsymbol{y}) = u_2(\boldsymbol{y}) - u_3(\boldsymbol{y})$, $\boldsymbol{y} \in S$ used in the representation (14) of $u_2(\boldsymbol{x})$. We now make the *Ansatz* that the frequency dependence of $\mu(\boldsymbol{y})$ is determined almost entirely by that of the illuminating wave $u_1(\boldsymbol{y})$. More precisely, we assume that

$$\mu(\boldsymbol{y}) = u_1(\boldsymbol{y}) \phi(\boldsymbol{y}), \tag{34}$$

* Here we are using the fact that $\int_{-\infty}^{+\infty} \overline{f(x)} g(x)\, dx = \int_{-\infty}^{+\infty} \overline{\hat{f}(y)} \hat{g}(y)\, dy$, where $f(x) = \exp(-i\lambda x^2/2)$ and $\hat{f}(y) = (i\lambda)^{-1/2} \exp(iy^2/2\lambda)$.

where $u_1(y)$ is given by (28) and $\phi(y)$ is a slowly varying function of y and λ. Substituting (34) into (15) and dividing through by $u_1(z)$, we find

$$1=\tfrac{1}{2}\phi(z)+\int_S e^{i\lambda|z-y|-i\lambda\omega\cdot(z-y)}\frac{\cos\beta}{4\pi|z-y|^2}\phi(y)dy, \tag{35}$$

where β is the angle between $\boldsymbol{\omega}$ and the outward normal $\boldsymbol{n}$.

When the surface is *convex* and λ is large, we can get an estimate for the integral in (35) by the stationary phase method. Apart from certain degenerate cases which we ignore here, the phase function $i\lambda(|\boldsymbol{z}-\boldsymbol{y}|-(\boldsymbol{z}-\boldsymbol{y})\cdot\boldsymbol{\omega})$ will be stationary only at a finite number of points, including those points $\boldsymbol{y}_i$ at which the vector $\boldsymbol{\omega}-\boldsymbol{\xi}_i$ is parallel to the normal $\boldsymbol{n}$. Here $\boldsymbol{\omega}$ is the illumination direction, and $\xi=(\boldsymbol{z}-\boldsymbol{y})/|\boldsymbol{z}-\boldsymbol{y}|$ is the direction of $\boldsymbol{z}-\boldsymbol{y}$. Thus y_i is a stationary point if the illumination is reflected off $\boldsymbol{y}_i$ to $\boldsymbol{z}$, according to the principles of geometric optics.

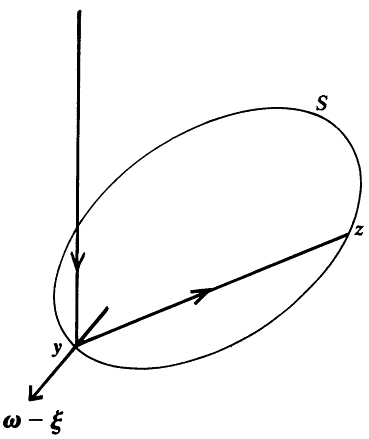


FIG. 1.

These points $\boldsymbol{y}_i$ all necessarily lie in the "lower half" of the surface where $\boldsymbol{\omega}\cdot\boldsymbol{n}>0$, i.e., "in the shadow". In addition, if $\boldsymbol{z}$ lies in the shadow, then there is normally one other stationary point $\boldsymbol{y}_0$, where $\boldsymbol{\omega}-\boldsymbol{\xi}_0=0$. This point lies on the "upper half" of the surface where $\boldsymbol{\omega}\cdot\boldsymbol{n}<0$, i.e., "in the light". A careful evaluation of (35) by stationary phase gives:

$$1=\tfrac{1}{2}\phi(z)-\sum_i\left(1-\frac{1}{i\lambda R_i}\right)e^{i\lambda R_i}\phi(y_i)+O\left(\frac{1}{\lambda^2R^2}\right) \tag{36}$$

or

$$1=\tfrac{1}{2}\phi(z)+\sum_i\left(1-\frac{1}{i\lambda R_i}\right)e^{i\lambda R_i}\phi(y_i)+\tfrac{1}{2}\left(1-\frac{1}{i\lambda R}\right)\phi(y_0)+O\left(\frac{1}{\lambda^2R^2}\right), \tag{37}$$

depending on whether $\boldsymbol{z}$ lies in the light (36), or in the shadow (37). Here $R_i=|\boldsymbol{z}-\boldsymbol{y}_i|-(\boldsymbol{z}-\boldsymbol{y}_i)\cdot\boldsymbol{\omega}$ is the phase factor of the stationary point $\boldsymbol{y}_i$, and R_0 vanishes at the exceptional point $\boldsymbol{y}_0$.

Now if λ is very large we can ignore terms in $1/\lambda$, except near the transition region from light to shadow where R is small. Within this approximation, we find that we can satisfy the equations (36) and (37) if we simply take

$$\phi(z)=\begin{cases}2, & z \text{ in the light}\\ 0, & z \text{ in the shadow.}\end{cases} \tag{38}$$

This choice of $\phi(z)$ is sometimes called the *physical optics approximation*, or Kirchhoff approximation, for ϕ; it is often chosen on the basis of the intuitively appealing physical argument that the induced current on S should be twice the illumination on the bright side and should vanish on the dark side. Here we see that it is valid up to terms of order $1/\lambda$, except near the transition region. There the situation is very complicated indeed [cf. **8**].

If we now substitute (38) into (14), we find the following representation for the scattered field:

$$u_2(x) = \int_{\omega \cdot n < 0} \frac{\partial G}{\partial n} 2u_1 dy. \tag{39}$$

Keeping only the leading term in the expansion (16) for G,

$$G(x, y) = \frac{e^{i\lambda r}}{4\pi r} e^{-i\lambda\omega \cdot y} + O\left(\frac{1}{\lambda r^2}\right) \tag{40}$$

we get

$$u_2(x) = \frac{e^{i\lambda r}}{4\pi r} F_0(\theta, \lambda) + O\left(\frac{1}{\lambda r^2}\right) \tag{41}$$

where

$$F_0(\theta, \lambda) = -2i\lambda \int_{\omega \cdot n < 0} e^{i\lambda(\omega - \theta)\cdot y}\, \theta \cdot n(y) dy \tag{42}$$

is the radiation pattern for $u_2(x)$.

Curvature and support for convex surfaces. Using the physical optics approximation for convex surfaces we can now evaluate the radiation pattern $F_0(\theta, \lambda)$ for large λ, using the stationary phase method and keeping only the leading terms. The result of a rather lengthy calculation is:

$$F_0(\theta, \lambda) = 2K^{-1/2}(\theta - \omega) \exp i\lambda P(\theta - \omega) + O(1/\lambda). \tag{43}$$

Here $K(\theta - \omega)$ is the *Gaussian curvature* of the surface S at the (unique) point whose normal is parallel to $\theta - \omega$ (the specular point) and $P(\theta - \omega)$ is the *support function* of S at that point, i.e., $P(\theta - \omega)$ is the distance from the origin to the plane tangent to S at that point.

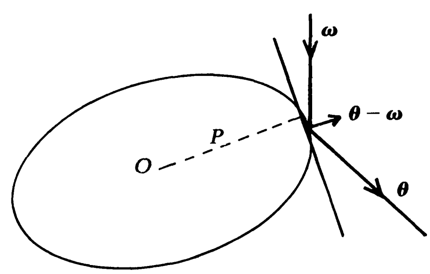


FIG. 2.

Thus we see that in this approximation the radiation pattern for a convex surface contains information about both the Gaussian curvature and the support function at each point on the illuminated side of the surface, and tells us nothing about the dark side. But if we now illuminate the surface from below instead of above, (i.e., replace ω by $-\omega$ in u_1), then the same argument will give the same data on the dark side.

We must expect to find trouble near the shadow boundary, however, since the physical optics approximation breaks down there; but if we stay outside of a thin strip along the shadow boundary

(i.e., if we restrict $\boldsymbol{\theta}$ away from a small cone in the forward $\boldsymbol{\omega}$ direction) then the physical optics and stationary phase approximations are all valid up to order $1/\lambda$ and indeed uniformly so.

To get information along the shadow boundary, uniformly valid to order $1/\lambda$, we have only to move the lamp again. A little thought shows that a total of five locations in all (i.e., above, below, and three along the equator 120° apart) will suffice.

Do the curvature and support functions suffice to determine the surface? Yes, they do. In fact, the curvature $K(\boldsymbol{n})$ alone suffices: every smooth convex surface is determined by its Gaussian curvature. This is in fact a famous problem in geometry, known as Minkowski's problem, which was solved in full generality only in 1953 by Nirenberg **[10, 19]**. The solution, unfortunately, is non-constructive and gives no useful information on how to reconstruct a surface from a knowledge of its curvature. A constructive solution still eludes us today,. in spite of numerous efforts.

The support function $P(\boldsymbol{n})$ also suffices: Every smooth convex surface is determined by its support function as the envelope of the family of planes with normal direction $\boldsymbol{n}$ and distance from the origin $P(\boldsymbol{n})$. Moreover, any finite number of these planes will give us a circumscribed approximating polyhedron, and the greater the number the better the approximation. Thus we are led to a simple constructive procedure for determining polyhedral approximations for S from the high frequency radiation pattern. Practically, however, these data are contained in the phase shifts, which are difficult to measure precisely.

The characteristic function. If we now introduce the characteristic function γ of B:

$$\gamma(x)=\begin{cases}1 & \boldsymbol{x}\in B\cup S\\ 0 & \boldsymbol{x}\in B'\end{cases} \tag{44}$$

and its Fourier transform

$$\Gamma(\lambda\boldsymbol{\xi})=(2\pi)^{-3/2}\int_3 e^{-i\lambda\boldsymbol{\xi}\cdot\boldsymbol{x}}\gamma(\boldsymbol{x})d\boldsymbol{x}, \tag{45}$$

then we can express the leading terms of the radiation pattern directly in terms of Γ. To this end, recall that

$$F_0(\boldsymbol{\theta},\lambda)=F(\boldsymbol{\theta},\boldsymbol{\omega},\lambda)=-2i\lambda\int_{S^+} e^{-i\lambda(\boldsymbol{\theta}-\boldsymbol{\omega})\boldsymbol{y}}\boldsymbol{\theta}\cdot\boldsymbol{n}d\boldsymbol{y}, \tag{46}$$

where S^+ is the illuminated portion of S, where $\boldsymbol{\omega}\cdot\boldsymbol{n}<0$. If we replace $\boldsymbol{\theta}$ by $-\boldsymbol{\theta}$ and $\boldsymbol{\omega}$ by $-\boldsymbol{\omega}$ (illuminate from below) then we get

$$F_0(-\boldsymbol{\theta},-\boldsymbol{\omega},\lambda)=2i\lambda\int_{S^-} e^{+i\lambda(\boldsymbol{\theta}-\boldsymbol{\omega})\boldsymbol{y}}\boldsymbol{\theta}\cdot\boldsymbol{n}d\boldsymbol{y}, \tag{47}$$

where now S^- is the shadow portion of S, where $\boldsymbol{\omega}\cdot\boldsymbol{n}>0$. Combining (46) and (47), we get

$$F_0(\boldsymbol{\theta},\boldsymbol{\omega},\lambda)+\overline{F_0(-\boldsymbol{\theta},-\boldsymbol{\omega},\lambda)}=-2i\lambda\int_S e^{-i\lambda(\boldsymbol{\theta}-\boldsymbol{\omega})\boldsymbol{y}}\boldsymbol{\theta}\cdot\boldsymbol{n}d\boldsymbol{y}. \tag{48}$$

We also know that $F_0(\boldsymbol{\theta},\boldsymbol{\omega},\lambda)=F_0(-\boldsymbol{\omega},-\boldsymbol{\theta},\lambda)$; i.e., that the radiation pattern is the same if we interchange the positions of the transmitter and the receiver. This, which is often called the *principle of reciprocity*, follows naturally from the symmetry of the Green's function G. Hence

$$\begin{aligned}F_0(\boldsymbol{\theta},\boldsymbol{\omega},\lambda)+\overline{F_0(-\boldsymbol{\theta},-\boldsymbol{\omega},\lambda)}&=F_0(-\boldsymbol{\omega},-\boldsymbol{\theta},\lambda)+\overline{F_0(\boldsymbol{\omega},\boldsymbol{\theta},\lambda)}\\&=2i\lambda\int_S e^{-i\lambda(\boldsymbol{\theta}-\boldsymbol{\omega})\cdot\boldsymbol{y}}\boldsymbol{\omega}\cdot\boldsymbol{n}d\boldsymbol{y}.\end{aligned} \tag{49}$$

Combining (48) and (49)

$$F_0(\boldsymbol{\theta}, \boldsymbol{\omega}, \lambda) + \overline{F_0(-\boldsymbol{\theta}, -\boldsymbol{\omega}, \lambda)} = -2i\lambda \int_S e^{-i\lambda(\boldsymbol{\theta}-\boldsymbol{\omega})\cdot y}(\boldsymbol{\theta} - \boldsymbol{\omega})\cdot \boldsymbol{n} dy \tag{50}$$
$$= 2\lambda^2(\boldsymbol{\theta} - \boldsymbol{\omega})^2 \int_B e^{-i\lambda(\boldsymbol{\theta}\cdot\boldsymbol{\omega})\cdot x} dx,$$

where we have used the divergence theorem to convert the surface integral to a volume integral. Comparing (50) with (45), we find

$$F_0(\boldsymbol{\theta}, \boldsymbol{\omega}, \lambda) + \overline{F_0(-\boldsymbol{\theta}, \boldsymbol{\omega}, \lambda)} = (2\pi)^{3/2} 2\lambda^2(\boldsymbol{\theta} - \boldsymbol{\omega})^2 \Gamma(\lambda(\boldsymbol{\theta} - \boldsymbol{\omega})). \tag{51}$$

Now we see that the combination $F_0(\boldsymbol{\theta}, \boldsymbol{\omega}, \lambda) + \overline{F_0(-\boldsymbol{\theta}, \boldsymbol{\omega}, \lambda)}$ is, in the physical optics approximation, a function of λ and $\boldsymbol{\theta} - \boldsymbol{\omega}$ only, and simply proportional to $\Gamma(\lambda(\boldsymbol{\theta} - \boldsymbol{\omega}))$. We expect this relation (51) to hold uniformly up to order $1/\lambda$, except near the shadow boundary, where $\boldsymbol{\theta} - \boldsymbol{\omega} = \boldsymbol{O}$.

We can apparently recover $\gamma(\boldsymbol{x})$ from $\Gamma(\lambda\boldsymbol{\xi}) = (\boldsymbol{\xi} - \boldsymbol{\theta} - \boldsymbol{\omega})$ via the Fourier transform. Some care is needed here, however, since Γ is given by the radiation pattern $F_0(\boldsymbol{\theta}, \lambda)$ only for large λ. We can, however take for $\Gamma(\lambda\boldsymbol{\xi})$ the leading term from $F_0(\theta, \lambda)$ for large λ, and then define $\Gamma(\lambda\boldsymbol{\xi})$ *in any smooth manner* for small λ. Then its Fourier transform will not be precisely $\gamma(\boldsymbol{x})$, but it will have the same singularity, namely, a unit jump at the surface S. This yields another reconstruction of S [cf. **2**].

Variations on this theme are also possible. For instance, suppose we can measure $F_0(\boldsymbol{\theta}, \lambda)$ for all large λ but for only one fixed $\boldsymbol{\theta}$. This will provide us with an estimate for $\Gamma(\lambda\boldsymbol{\xi})$ as above. Now for fixed $\boldsymbol{\xi} = \boldsymbol{\theta} - \boldsymbol{\omega}$ we have:

$$\Gamma(\lambda\boldsymbol{\xi}) = (2\pi)^{-3/2} \int_{\mathbf{R}^3} e^{i\lambda\boldsymbol{\xi}\cdot x} \gamma(\boldsymbol{x}) d\boldsymbol{x} = (2\pi)^{-1/2} \int_{-\infty}^{+\infty} e^{i\lambda t} A(t) dt, \tag{52}$$

where $A(t)$ is the *cross-sectional area function*

$$A(t) = (2\pi)^{-1} \int_{x\cdot\boldsymbol{\xi}=t} \gamma(\boldsymbol{x}) d\boldsymbol{x}. \tag{53}$$

Thus $\Gamma(\lambda\boldsymbol{\xi})$ for fixed $\boldsymbol{\xi}$ is the Fourier Transform of the area function giving the areas of the cross sections of B in planes with normal $\boldsymbol{\xi}$.

If we now vary $\boldsymbol{\xi}$ over any two distinct vertical planes, then we have a handful of data describing B. Is it enough to determine B? Yes, it is; every smooth convex surface is determined by its cross sectional areas taken over any two distinct circles of directions. This is another famous problem of surface geometry, known as Radon's problem, whose solution can be compactly expressed in terms of Radon transforms [cf. 7].

Finally we note that we can express $\Gamma(\lambda\boldsymbol{\xi})$ in terms of the leading terms of the spherical moments of u_2:

$$2\lambda^2(\boldsymbol{\theta} - \boldsymbol{\omega})^2 \Gamma(\lambda(\boldsymbol{\theta} - \boldsymbol{\omega})) = F_0(\boldsymbol{\theta}, \boldsymbol{\omega}, \lambda) + \overline{F_0(-\boldsymbol{\theta}, -\boldsymbol{\omega}, \lambda)} \tag{54}$$
$$= \sum_{n=0}^{\infty} \sum_{m=-n}^{m} i^{-n-1} C_{mn}(f_{mn}(\infty) + f_{mn}(-\infty)) Y_{mn}(\boldsymbol{\theta}) + O(1/\lambda).$$

Thus

$$f_{mn}(\infty) + \overline{f(-\infty)} = g_{mn} + O(1/\lambda), \tag{55}$$

where

$$g_{mn} = C_{mn} \int 2\lambda^2(\boldsymbol{\theta} - \boldsymbol{\omega})^2 \Gamma(\lambda(\boldsymbol{\theta} - \boldsymbol{\omega})) \overline{Y_{mn}(\boldsymbol{\theta})} d\boldsymbol{\theta}. \tag{56}$$

But

$$\lambda^2\xi^2\Gamma(\lambda\xi)=(2\pi)^{-3/2}\int_B \nabla^2 e^{i\lambda\xi\cdot y}dy=(2\pi)^{-3/2}\int_S \frac{\partial}{\partial n}e^{i\lambda\xi\cdot y}dy \tag{57}$$

and

$$e^{i\lambda\theta\cdot y}=\sum C_{mn}Y_{mn}(\theta)j_n(\lambda\,|\,y\,|). \tag{58}$$

Combining (56), (57), and (58) ($\xi=\theta-\omega$) we finally find

$$g_{mn}=(2\pi)^{-3/2}\int_S \bar{Y}_{mn}(\theta)\frac{\partial}{\partial n}(2j_n(\lambda\,|\,y\,|)e^{i\lambda\omega\cdot y})dy. \tag{59}$$

It follows that the surface S can also be reconstructed from the moments g_{mn} by solving the moment problem (59). In particular, the first few moments give a crude global approximation to the surface.

Summary. We have seen that, in principle, the shape of a smooth connected compact surface in $\mathbb{R}^3$ is determined by the eigenfunctions of the exterior Dirichlet problem, and indeed, by their radiation patterns at large distances from the surface. It suffices to know these radiation patterns over a small band of incident directions or frequencies. In practice, however, the reconstruction of the surface involves an analytic continuation process which is unstable against small errors in data and inaccessible to approximation techniques.

At low frequencies, however, we can reconstruct the surface from its static moments via the static potential. And at high frequencies we can reconstruct a convex surface from its Gaussian curvature, its support function, or its characteristic function, using the physical optics approximation. The support function is particularly suitable, since it allows us to construct a polyhedral approximation to the surface from a finite number of phase measurements. But better approximation techniques are needed, and the questions of non-convex and non-smooth surfaces have hardly been touched.

Added in Proof: The author is indebted to Prof. Joseph Keller of NYU for his lively conversations and correspondence on the general subject of inverse problems, for several helpful suggestions, and for a preprint of his recent paper **[20]**.

References

1. N. N. Bojarski, Inverse scattering, Syracuse Univ. Res. Corp. Special Projects Lab. Reports, AD845, #122–127 (1968).

2. ———, Electromagnetic inverse scattering, NASA Tech. Memorandum TMX–62, #150 (1972).

3. M. Kac, Can one hear the shape of a drum? this MONTHLY, 73 (1966 II) 1–23.

4. J. Keller, The inverse scattering problem in geometric optics, IRE Trans. Antennas and Propagation, AP–7 (1959) 146–149.

5. R. E. Kleinman, The Dirichlet problem for the Helmholtz equation, Arch. Rational Mech. Anal., 18 (1965) 205–229.

6. P. Lax and R. Phillips, Scattering Theory, Academic Press, New York, 1967.

7. R. Lewis, Physical optics inverse diffraction, IRE Trans. Antennas and Propagation, AP–17 (1969) 308–314.

8. D. Ludwig, Boundary layers in the field scattered by a convex object, Comm. Pure Appl. Math., 9 (1965) 115–134.

9. P. Morse and H. Feshbach, Methods of Theoretical Physics, McGraw–Hill, New York, 1953.

10. L. Nirenberg, The Weyl and Minkowski problems, Comm. Pure Appl. Math., 6 (1953) 337–399.

11. A. Sommerfeld, Die Greensche Funktion der Schwingungsgleichung, Jber. Deutsch. Math.-Verein, 21 (1912) 309–353.

12. ———, Partial Differential Equations in Physics, Academic Press, New York, 1949.

13. J. A. Stratton, Electromagnetic Theory, McGraw-Hill, New York, 1941.

14. V. Twersky, Inversion of radiation data in biophysics, NASA Tech. Memorandum TMX–62, #150 (1972).

15. V. Weston, J. Bowman, and E. Ar, On the electromagnetic inverse problem, Arch. Rational Mech. Anal., 31 (1968) 199–213.

16. C. Wilcox, A generalization of theorems of Rellich and Atkinson, Proc. Amer. Math. Soc., 7 (1956) 271–276.

17. ———, Analytic continuation of the resolvent kernel in scattering theory, Edinburgh Conf. Partial Differential Equations, 1967.

18. H. Weyl, Die natürliche Randwertaufgaben im Aussenraum für Strahlungsfelder, Math. Zeit., 56 (1952) 104–119.

Added in Proof:

19. Hans Lewy, On differential geometry in the large, I, Trans. Amer. Math. Soc., 43 (1938) 258–270.

20. J. Keller, Inverse problems, this MONTHLY, 83 (1976) 107–118.

21. O. D. Kellogg, Foundations of Potential Theory, Springer, Berlin, 1929.

DEPARTMENT OF MATHEMATICS, DARTMOUTH COLLEGE, HANOVER, NH 03755.

THE EXISTENCE OF LIMITS FOR SOLUTIONS OF SINGULAR DIFFERENTIAL INEQUALITIES

HARRY GINGOLD AND RAY REDHEFFER*

1. Introduction. If a real-valued function y of the real variable t satisfies a suitable differential inequality on an interval, it is often possible to deduce existence of $\lim y(t)$ as t approaches either end of the interval. Suppose, for example, that

$$|y'(t)| \leqq M \quad \text{for} \quad -\varepsilon \leqq t < 0, \tag{1}$$

where M and ε are positive constants. Then the function $y(t) - Mt$ is decreasing as $t \to 0-$, hence is bounded above. This shows that $y(t) + Mt$ is also bounded above. Since the latter function is increasing, it has a limit, L. Therefore $y(t)$ has the same limit L.

This well-known consequence of (1) extends to inequalities of higher order. If

$$|y^{(n)}(t)| \leqq M \quad \text{for} \quad -\varepsilon \leqq t < 0 \qquad (n \geqq 2), \tag{2}$$

then (1) applied to $y^{(n)}$ instead of y' shows that $y^{(n-1)}(0-)$ exists. Hence $y^{(n-1)}(t)$ is bounded near $0-$, and we can apply (1) again. The process continues in an obvious manner, and establishes that

$$y(0-), y'(0-), \ldots, y^{(n-1)}(0-) \quad \text{exist.} \tag{3}$$

Our objective is to formulate results similar to those above for tolerably general classes of differential inequalities; actually the form of the inequalities could be further generalized, as will be obvious from the proofs. The most important aspect of our analysis is the fact that *the inequalities are allowed to be singular*, in the sense that the leading coefficient can vanish on an arbitrary set of points. This singular behavior makes it difficult to apply the more familiar theorems pertaining to differential equations, and we have not happened to come across any previous results which imply the results given here. Nevertheless, the methods are elementary, and seem to us to be well suited for classroom presentation.

2. Notation. For ease of exposition we agree that m, M, ε, a, b denote constants, with $\varepsilon > 0$. The letters p, q, y, q_j denote real-valued functions defined on $[-\varepsilon, 0)$, and all hypotheses are understood to hold in this interval for some ε. A condition holds "near $0-$" if it holds on $-\delta \leqq t < 0$ for some $\delta > 0$.

* Under auspices of the U.S. Special Program, Alexander von Humboldt-Stiftung

Here δ may agree with the ε of the hypothesis, but need not do so. As usual, $y(0-)$ denotes the limit of $y(t)$ as $t \to 0$ through negative values; this notation was already used in (3).

3. A simple limit theorem. Our first result seems to require a differential equation, but actually pertains to the inequality $m < py' + y < M$.

THEOREM 1. *Let $py' + y = q$ where $p \geqq 0$ and $m < q < M$. Then one of the following holds: $y(0-)$ exists, or $m < y(t) < M$ near $0-$.*

For proof, write $py' = q - y$. If $y \geqq M$ for all t near $0-$ then $y' < 0$, hence y is decreasing and bounded below, so has a limit. Similarly, if $y \leqq m$ for all t near $0-$ then y has a limit.

Suppose next that $m < y(t_0) < M$ for some t_0 near $0-$. Then $y' < 0$ when $y = M$ and hence, y does not escape from $m < y < M$ at the top boundary $y = M$. Similarly, y does not escape from the bottom. This completes the proof. The choice $p = t$, $y = 1/t$ shows that $p \geqq 0$ is essential.

THEOREM 2. *Let $py' + y = q$ near $0-$, with $p \geqq 0$. Then $y(0-)$ exists, or else, as $t \to 0-$,*

$$\liminf q(t) \leqq \liminf y(t) \leqq \limsup y(t) \leqq \limsup q(t).$$

This follows by applying Theorem 1 on $(-\varepsilon_i, 0)$ with $\varepsilon_i > 0$, $\varepsilon_i \to 0$. As a special case, we see that if q is bounded then y is bounded, and if $q(0-)$ exists then $y(0-)$ exists. The latter assertion is generalized in Theorem 3.

THEOREM 3. *If $py^{(n+1)} + y^{(n)} = q$ where $p \geqq 0$ and where $q(0-)$ exists, then*

$$py^{(n+1)}, y^{(n)}, \ldots, y', y$$

all have limits at $0-$.

This follows by applying Theorem 2 to $\tilde{y} = y^{(n)}$ and then using (2) $\Rightarrow$ (3). A more general form of Theorem 3 (which is also harder to prove) is given in the following section.

4. Inequalities containing a norm. With $\|y\|$ defined by

$$\|y\|(t) = \max |y(s)| \quad \text{for} \quad -\varepsilon \leqq s \leqq t < 0,$$

we shall establish the following:

THEOREM 4. *Let $|py' + y| \leqq a + b\|y\|$ where $p \geqq 0$ and $b < 1$. Then y is bounded.*

For proof, choose M so that $M > |y(-\varepsilon)|$ and $M > a + bM$. It will be shown that $|y(t)| < M$. If not, let t be the first value beyond $-\varepsilon$ where $|y(t)| = M$. Without loss of generality, $y(t) = M$. Then the hypothesis at t implies

$$|p(t)y'(t) + M| \leqq a + bM < M.$$

This shows that $y'(t) < 0$. Hence $y(t-h) > y(t) = M$ for some $h > 0$, which contradicts the fact that t was the first point where $y = M$.

The norm $\| \quad \|$ can be used to estimate a function in terms of its derivative. For example, if y' is continuous and $-\varepsilon \leqq t < 0$, then

$$|y(t)| = \left| \int_{-\varepsilon}^{t} y'(s)ds + y(-\varepsilon) \right| \leqq \varepsilon \|y'\|(t) + |y(-\varepsilon)|. \tag{4}$$

By repeated integration we can estimate $y^{(j)}$ similarly, in terms of $y^{(n)}$, and addition gives an inequality of form

$$\|y\| + \|y'\| + \cdots + \|y^{(n-1)}\| \leqq \alpha + \beta\varepsilon \, \|y^{(n)}\|, \tag{5}$$

where α and β are constant. This result is used together with Theorem 4 in the next section.

5. Application to higher-order inequalities. The results of §4 lead to:

THEOREM 5. *Let* $|py^{(n+1)}+y^{(n)}|\leqq a(1+|y|+\cdots+|y^{(n-1)}|)$ *where* $p\geqq 0$. *Then* $y^{(n)}$ *is bounded near* $0-$, *and hence* $y, y', \ldots, y^{(n-1)}$ *all have limits at* $0-$.

We shall establish this under the weaker hypothesis

$$|py^{(n+1)}+y^{(n)}|\leqq a(1+\|y\|+\cdots+\|y^{(n-1)}\|). \tag{6}$$

First decrease ε, if necessary, so that $\varepsilon\beta a<1$ in (5). With

$$A=a\alpha+a, \qquad B=\varepsilon\beta a, \qquad Y=y^{(n)},$$

equations (5) and (6) give

$$|pY'+Y|\leqq A+B\|Y\|.$$

Theorem 4 shows that Y is bounded, and the remaining assertions follow from (2) $\Rightarrow$ (3).

An interesting consequence of Theorem 5 is the following improved form of Theorem 3:

THEOREM 6. *Let* $q_j(0-)$ *exist for* $j=0,1,\ldots,n$, *let* $p\geqq 0$, *and let*

$$py^{(n+1)}+y^{(n)}+q_{n-1}y^{(n-1)}+\cdots+q_1y'+q_0y=q_n.$$

Then $py^{(n+1)}$, $y^{(n)},\ldots,y'$, y *all have limits at* $0-$.

For proof, write the differential equation as $py^{(n+1)}+y^{(n)}=q$ where, by definition,

$$q=q_n-q_0y-q_1y'-\cdots-q_{n-1}y^{(n-1)}.$$

Since $q_j(0-)$ exist, $|q_j|$ are bounded near $0-$, and we can diminish ε, if necessary, so that $|q_j|$ are bounded on $[-\varepsilon,0)$. If a is a common bound for all $|q_j|$ then

$$|q|\leqq a(1+|y|+\cdots+|y^{(n-1)}|).$$

Theorem 5 shows that $y_j(0-)$ exist for $j=0,1,\ldots,n-1$, hence $q(0-)$ exists also, and the conclusion follows from Theorem 3.

To formulate a similar result for nonlinear equations, let $f(t,\mathbf{z})$ be a function from $[-\varepsilon,0]\times\mathbb{R}^n$ to $\mathbb{R}$. We say that $f(t,\mathbf{z})$ is Lipschitzian at $\mathbf{z}=\mathbf{0}$ if

$$|f(t,\mathbf{z})-f(t,\mathbf{0})|\leqq\beta|\mathbf{z}|, \qquad -\varepsilon\leqq t<0, \tag{7}$$

for some constant β. Since β can be increased at will, the choice of norm $|\mathbf{z}|$ does not matter.

THEOREM 7. *Let* $p\geqq 0$, *let* $f(t,\mathbf{z})$ *be Lipschitzian at* $\mathbf{z}=\mathbf{0}$, *and let*

$$py^{(n+1)}+y^{(n)}=f(t,y,y',\ldots,y^{(n-1)}).$$

Suppose further that f *is continuous at the boundary points of its domain where* $t=0$. *Then* $py^{(n+1)}, y^{(n)},\ldots,y', y$ *all have limits at* $0-$.

For proof, write the differential equation in the form

$$py^{(n+1)}+y^{(n)}=[f(t,\mathbf{y})-f(t,\mathbf{0})]+f(t,\mathbf{0})$$

and estimate the bracket by (7). Under the hypothesis that $f(t,\mathbf{0})$ is bounded near $0-$, we can apply Theorem 5. Under the stronger hypothesis that $f(t,\mathbf{z})$ is continuous at $(0,\mathbf{z})$ we can then apply Theorem 3, and the result follows.

6. Monotone behavior. We shall establish the following:

THEOREM 8. *Let* $|py'+y|\leqq a+b\|y\|$ *where* $b<1$. *Suppose further that the set of zeros of* y' *has* 0 *as a limit point. Then* y *is bounded.*

For proof, let $y'(t_j)=0$ where $t_j \to 0$. Let M_j be the maximum of $y(t)$ on $[-\varepsilon, t_j]$ and let M be the larger of the two numbers $|y(-\varepsilon)|$, $a/(1-b)$. We shall show $M_j \leqq M$; then $|y| \leqq M$ follows from $t_j \to 0-$. Suppose, then, that $M_j > M$, and let t be the first point beyond $-\varepsilon$ where $|y(t)| = M_j$. Then $y'(t)=0$, $\|y\|(t)=M_j$, and the differential inequality leads to a contradiction.

Theorem 8 indicates that if $|y|$ is unbounded then $y'(t)$ cannot change sign near $0-$, and hence, y is strictly monotone near $0-$. Another theorem pertaining to monotone behavior is given next.

THEOREM 9. *Let $|py''+qy'+y| \leqq a+b\|y\|$ where $p \leqq 0$ and $b<1$. Then either y is bounded, or y is strictly monotone near $0-$.*

For the benefit of readers who may want to check their mastery of the techniques presented here, we omit the easy proof.

Theorems 8 and 9 can be applied to equations of higher order, as is evident from the discussion of section 5. The following rather curious result, for example, is implied by Theorem 9:

THEOREM 10. *Let $p \leqq 0$ and let*

$$|py^{(n+2)}+qy^{(n+1)}+y^{(n)}| \leqq a(1+|y|+\cdots+|y^{(n-1)}|.$$

Suppose further that $y^{(n)}$ does not diverge monotonely to $+\infty$ or to $-\infty$ at $0-$. Then $y, y', \ldots, y^{(n-1)}$ all have limits at $0-$, and $py^{(n+2)}+qy^{(n+1)}$ is bounded.

Theorem 10 follows from Theorem 9 in the same way that Theorem 5 followed from Theorem 4.

7. Change of variable. By the substitution $t=-1/x$ the interval $[-\varepsilon, 0)$ is transformed to $[1/\varepsilon, \infty)$. Hence, the foregoing theorems have analogs in which the condition $x \to \infty$ replaces $t \to 0-$. These results will not be formulated here, but we mention that they provide an interesting supplement to some of those in the following references.

References

1. Lamberto Cesari, Asymptotic Behavior and Stability Problems in Ordinary Differential Equations, Springer-Verlag, New York, 1963, Second Edition, pp. 36–42.

2. Earl A. Coddington and Norman Levinson, Theory of Ordinary Differential Equations, McGraw-Hill, New York, 1955, p. 61, 99.

3. Philip Hartman, Ordinary Differential Equations, Wiley, New York, 1964, p. 273–277, 350, 501, 508–510.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF UTAH, SALT LAKE CITY, UT 84112.
MATHEMATISCHES INSTITUT I, UNIVERSITÄT KARLSRUHE, ENGLERSTRASSE 2, D-75 KARLSRUHE, GERMANY.

THE MOORE METHOD

F. BURTON JONES

While one cannot say that the "Moore Method" of teaching mathematics has gained wide-spread acceptance in college and university circles, it has been and is being successfully used by enough people to attract attention — even outside academia. Furthermore, there is considerable interest and curiosity about what it is and how it works. Mainly, for this reason, I am going to describe my own experiences in Moore's classes and in using the method with my own students, both graduate and undergraduate.

A bit of history. While he was still a graduate student at the University of Chicago (1903–05), R. L. Moore conceived the basic ideas that led eventually to his rather radical method of teaching. With his quick mind and restless spirit he found the lecture method rather boring — in fact, mind dulling. To liven up a lecture he would run a race with his professor by seeing if he could discover the proof of an announced theorem before the lecturer had finished his presentation. Quite frequently he won the race. But in any case, he felt that he was better off from having made the attempt. So if one could get students to prove the theorems for themselves, not only would they have a deeper and longer lasting understanding, but somehow their ability and interest would be strengthened. If the theorems were too difficult, then they would have to be broken down into easier lemmas.

The more he thought about having the students discover for themselves the mainstream of a subject, the more he became convinced that it not only could be made to work but that it would also be attractive to students. He spoke of this plan to one of his professors (possibly Veblen) who abruptly replied: "Ha, let the students do the work!" But E. H. Moore's reaction was more thoughtful — perhaps the approach did have merit.

As a beginning instructor it was not easy for Moore to find a department where he had sufficient freedom to give the idea a really good try. But when he did, he began (at The University of Pennsylvania) to have success, especially in the Foundations of Geometry. Here was a fresh, relatively new area where Moore had himself tested the difficulty of some of the theorems. In the years following his appointment at The University of Texas, he expanded the use of his limited lecture method to all of his classes: Calculus, Advanced Calculus, Measure Theory, Metric Density, etc., as well as Foundations of Geometry and (at the graduate level) Topology. The results were, from the standpoint of research productivity after the Ph.D., really phenomenal. Of the Ph.D.'s produced in the U.S. and Canada during the period 1915–1954, 25% of those from Texas were among the top 15% in the nation in productivity; 5% of those from U.C. (Berkeley) were in the top 15%; 8% of those from Chicago were in the top 15%; 16% of those from Harvard were in the top 15%; 20% of those from Princeton were in the top 15%. (A Survey of Research Potential and Training in the Mathematical Sciences, The University of Chicago, March 15, 1957; *The Albert Report.*) And it should be no surprise that high quality was associated with high productivity. Included in this group from Texas are Moore's students R. L. Wilder, G. T. Whyburn, J. H. Roberts, R. H. Bing, E. E. Moise, R. D. Anderson, and Mary Ellen Rudin (in chronological order).

What Moore did. Moore would begin his graduate course in topology by carefully selecting the members of the class. If a student had already studied topology elsewhere or had read too much, he would exclude him (in some cases, he would run a separate class for such students). The idea was to have a class as homogeneously ignorant (topologically) as possible. He would usually caution the group not to read topology but simply to use their own ability. Plainly he wanted the competition to be as fair as possible, for competition was one of the driving forces. (For the foundations of geometry he made no attempt to select the students because all of them, young and old, high school teachers or not, were uniformly ignorant of the Hilbert–Veblen–Moore axiomatic approach to the subject.)

Having selected the class he would tell them briefly his view of the axiomatic method: there were certain undefined terms (e.g., "point" and "region") which had meaning restricted (or controlled) by the axioms (e.g., a region is a point set). He would then state the axioms that the class was to start with (Axioms 0 and 1 of his book: *Foundations of Point Set Theory*, omitting part (4) of Axiom 1). An example or two of situations where the axioms could be said to apply (e.g., the plane or Hilbert space) would be given. He would sometimes give a different definition of region for a familiar space (e.g., Euclidean 3-space) to give some intuitive feeling for the meaning of an "undefined term" in the axiomatic system. Of course, this was part of his own personal philosophy and he considered it part of the motivation of the subject.

After stating the axioms and giving motivating examples to illustrate their meaning he would then state some definitions and theorems. He simply read them from his book as the students copied them

down. He would then instruct the class to find proofs of their own and also to construct examples to show that the hypotheses of the theorems could not be weakened, omitted, or partially omitted.

When the class returned for the next meeting he would call on some student to prove Theorem 1. After he became familiar with the abilities of the class members, he would call on them in reverse order and in this way give the more unsuccessful students first chance when they *did* get a proof. He wasn't inflexible in this procedure but it was clear that he preferred it.

When a student stated that he could prove Theorem x, he was asked to go to the blackboard and present his proof. Then the other students, especially those who hadn't been able to discover a proof, would make sure that the proof presented was correct and convincing. Moore sternly prevented heckling. This was seldom necessary because the whole atmosphere was one of a serious community effort to understand the argument.

When a flaw appeared in a "proof" everyone would patiently wait for the student at the board to "patch it up." If he could not, he would sit down. Moore would then ask the next student to try or if he thought the difficulty encountered was sufficiently interesting, he would save that theorem until next time and go on to the next unproved theorem (starting again at the bottom of the class).

Occasionally theorems got left over indefinitely but nearly all of these would be proved in some subsequent year.

Quite frequently when a flaw would appear in a proof everyone would spend some time (possibly in class) trying to get an example to show that it couldn't be "patched up," i.e., a counterexample to the argument (even though the theorem might be correct). This kind of experience is seldom encountered in courses or in any place outside of one's own research work. Yet this kind of activity is vitally necessary for the research worker.

Occasionally an improvement on one of Moore's theorems would be proved. Moore would then refer to that theorem with the student's name. (In the revision of his book a number of names like this appear in the text and in some cases he made remarks concerning the origin of certain proofs and concepts in the appendix.) The improvement might not be very significant but the encouragement given by the "public recognition" was considerable. At the same time, mistakes were not discouraged. However, on certain theorems (the arcwise connectivity theorem, for example) knowing from experience that mistakes were likely, Moore would insist that the student have the proof written out before presenting it. This tended to reduce the number of false starts (and also gave the student some useful writing practice).

Difficulties and drawbacks. Probably the most obvious difficulty is the one of class uniformity. If the competition is not reasonably on an equal footing, then one of the basic drives has been weakened. Obviously, in contrast to the 1920's, little can be done about this now when almost all undergraduates have had some exposure to topology and topological ideas. G. T. Whyburn spoke to me more than once about how to handle this problem. My solution (or partial solution) is to begin general topology with only the most basic axiom: there exists a non-empty set S of elements called points and a collection G of subsets of S such that G covers S. Calling the elements of G "regions" and using the usual definitions of "limit point," "closed," and "open," some of the usual propositions about these notions are false but when carelessly formulated the usual arguments seem to work. This forces a kind of uniformity on the class. (I continue this introductory course by supposing that the space is semi-metric, i.e., metric without the triangle inequality, and everything returns to normal — derived sets are closed, the union and intersection of closed sets are closed, etc.) However, I do not find the lack of uniformity to be a severe handicap because the class becomes more homogeneous as to background as time goes on.

There is a problem of what to do with students who are too timid to present their proofs at the board. I generally try to draw them into the discussions, offer to do the writing on the board for them while they stay seated, and eventually after six months or so they get up without realizing what they are doing — especially if a subtle argument gets rather heated. But when there are students who don't

present theorems or counterexamples in class, I simply depend upon the final examination to determine the course grade. For those not actively participating, the course becomes a lecture course (the lectures lacking somewhat in polish).

There is also the student whose proof gets "shot down." People generally tend to be embarrassed by mistakes — especially public mistakes — and care must be taken not to make them feel that criticism is ridicule. As working mathematicians we have become so accustomed to making mistakes that we are inclined to forget the pain experienced by the young.

Then there is the problem of *reading*. After a few months of working out their own proofs it becomes quite difficult to get students to read mathematics. They would rather do it themselves. Nevertheless, reading is necessary if one is to become educated (especially in the brief span of graduate school). Moore used to complain that when he wanted (finally) the students to read, they couldn't. After some trial and error I have found the following technique to be effective. Have the students in the beginning graduate general topology not read until Christmas and then buy themselves a copy of Kelley. They can then use Kelley for bedtime reading to prepare for the final examination (and the qualifying examination — I make them the same) which will be mostly (70%) on Kelley. I never discuss Kelley in class. The students are supposed to do this reading unassisted. Class continues as before with the students working out their own proofs for by then we are going in a direction which overlaps Kelley very little.

By far the most difficult aspect of the method is patience. The instructor must not help — must not point out the "obvious." We all know how difficult even the obvious is before it becomes obvious. The instructor must simply be willing to wait for the students' mental chemistry to work. It helps if the instructor feels rewarded when the student *does* finally see how to put together a few ideas correctly.

And finally there is the problem of what to do when "no one has anything." One can, of course, start a new topic, motivate it, and get the students to construct some examples to fit, or state some new definitions and theorems. But this kind of thing cannot be done very often for there must not be too many unsettled problems or the student will become distracted and unable to concentrate his attack. I have several topics which can become useful side issues: (1) problems about set theory, well-ordering and cardinality can be taken up and worked out by the students on the spur of the moment, (2) how some of the theory simplifies if one assumes the space to be metric, (3) the history of some of the ideas and the personalities of some of the people involved, etc. But generally I make up some questions involving the application of theorems already proved (or even those yet to be proved) which can be settled in ten or fifteen minutes each.

At such times one can introduce the beginnings of notions, usually in connection with examples, that will be useful on theorems to come. That is, one can somewhat randomly (and out of context) put into the subconscious minds of the students pictures, ideas and notions which will resurface weeks (or even months) later in a proof. There should be no hint as to what the ideas are really for and in fact the student when he later uses one of the notions will have the feeling that he discovered it himself. In this way, as well as the actual statement of applicable lemmas, the proofs of quite difficult theorems can be made accessible.

Where can the method be used? One frequently hears people say: "Yes, it may work in topology but not in ...." I will discuss briefly certain areas where I have found it to work very well.

One of the most rewarding is group theory. Here again one may begin with rather simple axioms and the central theory can be broken down into a rather nice sequence of interesting but easy theorems. With two different classes I have used Speiser's *Die Theorie der Gruppen* with modifications (numerous illustrative examples should be presented by both the students and the instructor and Theorem 40 (Frobenius) is too difficult without the introduction of a lemma).

As previously mentioned the axiomatic development of the Hilbert–Veblen–Moore Foundations of (plane) Geometry works extremely well. There are proofs in this subject for students of widely varying abilities (the "mid-point" theorem: If ab is an interval, it contains a point c such that ac is congruent to cb, is too difficult without the Dedekind Cut Axiom).

An introduction to Hilbert Space works very well from von Neumann's 1928 paper (see M. H. Stone's colloquium publication — some of the lemmas may be omitted and left to the students to discover). The start of this theory makes a nice six to ten week course. I once told von Neumann about doing this and he remarked that it would require a very good class. The class was good and being a summer course, the students were already experienced in "proving things for themselves." However, once the ideas are stated and the sequence of theorems laid out, the rest is much, much easier than the situation von Neumann faced.

These are areas where a set of axioms can be used for a beginning. What about other areas ? I have personally used the method in courses in real variable theory and in complex variable theory. H. S. Wall and some of his students, MacNerney and Porcelli, in particular, have used the method (with obvious success) in various courses in analysis. Some areas require quite some effort in formulating a good sequence of theorems. MacNerney worked two years to develop a sequence (in Complex Variable Theory) which would yield the Cauchy Integral Theorem in one quarter (one semester is better).

For the topology of the line (and the plane) I have found the non-axiomatic approach to be more successful (for beginning college and high school students). I spend some time being certain that the class is clear on the properties of the natural numbers including order, well-order and induction. A "point" is defined to be a function from the natural numbers to the natural numbers. Then giving the set of all such functions the lexicographic order and using the open interval topology (except at the constant function 1 where the correct half-open interval is necessary), one has an interesting beginning of a development that quickly gives all the topological properties of the non-negative real numbers. The "sum" and "product" of points given by the usual definitions for functions is discontinuous. Hence the discovery of a perfect set which contains no interval presents the student with something like the situation that Cantor encountered.

Of course, there is also a simple axiomatic approach to the topology of the line (which omits arithmetic) and Burgess has used it quite successfully.

Moore often gave a summer course in metric density in the plane and the line where the beginning was non-axiomatic. In fact, a cursory knowledge of Lebesgue measure was required.

Who can use the method? I have already mentioned that the instructor must possess "patience." But I think it is more than that. It must be "patience" that is born of the conviction that training a student to do research is important — even more important than conveying knowledge; that trying to develop a student's mathematical ability to the limit of that ability is important.

The instructor should suppress his own urge to get into the act. Even an ugly proof from one of the students should please the instructor. Only on rare occasions (possibly "never" is the better policy) should he show an elegant proof. The student can learn "elegance" from his reading later. Moore once did this to me and his elegant argument drove my ugly one out of my mind and I have wished many times later that I could recall it where Moore's technique would not work. Maybe my ugly one would have. This is not to say that wording and clarity of thought should not be discussed. But it should be done on the student's own argument.

The instructor should work out the sequence of theorems keeping in mind the general needs and abilities of a group as the course proceeds so that extra lemmas may be introduced (for a weak group) or some lemmas may be omitted (for a strong group). I even like to leave the general direction of the course flexible so as to accommodate the interests of the group as I become better acquainted with the various individuals. The main thing, as often expressed by W. M. Whyburn, is to "give them something they can do."

Wilder has expressed the opinion that Moore was successful in using the method because the students were proving Moore's theorems while the development of the theory was still hot. I expect this may have been a factor in the obvious enthusiasm in the Moore School of the 1920's. This was absent later (say in the 1950's or a bit earlier) and Moore was just as successful. Clearly to be

successful, the instructor should give the student puzzle after puzzle that the student cannot resist. His appetite (and ability) increases with every solution.

Techniques. In teaching General Topology several people have followed Moore's technique of simply assigning the class the sequence of theorems as they occur in Moore's book. Most people, however, have used some variant of the method more suitable to their own ideas of what develops the student. For example, Bing has some student in the class act as secretary and write up the notes for the whole class. By passing this chore around, some of the students get some practice at writing. (And it's a good idea for other reasons.) I like to state false propositions (just as if they were true) for the students to prove. And quite frequently I state a number of definitions and ask the students to formulate some theorems using them. I feel that examples (and counterexamples) are very important for both understanding and motivation. In particular, one's intuition is aided by examples of spaces that do not satisfy the axioms as much as by examples that do: for instance, (in my approach to General Topology) topological spaces (even compact and Hausdorff) that are not semi-metric and semi-metric spaces which are not metric.

And it's nice to have a sequence of theorems which are useful but which can be proved by anybody. Elementary properties of connected point sets can be formulated into a sequence of this sort.

It is a good plan to encourage students to change a theorem until they can prove it: weaken the conclusion or strengthen the hypothesis or both. This helps to avoid frustration and is good practice.

Final remarks. The instructor should convey confidence, especially in the beginning. The student should soon learn that some things he can do quickly but others may take effort and time. Given six months practice, a student who had never thought of a proof in his life and didn't know how to start, may develop to the point where he can settle almost anything you propose. I find this very rewarding and satisfying. It happens often enough to keep one's enthusiasm for teaching vitally alive.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, RIVERSIDE, CA 92502.

MATHEMATICAL NOTES

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Laboratoire Calcul des Probabilités, Université de Paris, T.56, 4 Place Jussieu, 75-230 Paris, France.

A CHARACTERIZATION OF INTEGERS

MAURICE MIGNOTTE

Abstract. A necessary and sufficient condition for a real number $\theta > 1$ to be a rational integer is to verify

$$\|\theta^n\| \leqq ((\theta+1)(\theta+2))^{-1} \quad \text{for} \quad n = 1, 2, \ldots,$$

where $\|\ \|$ denotes the distance to the nearest integer.

1. Introduction. In this paper θ is a real number > 1. For real x we put $\|x\| = \min_{n \in \mathbb{Z}} |x - n|$; this is the distance to the nearest integer.

We are interested in elementary results about the behaviour of $\|\lambda\theta^n\|$. Our first result shows (see Corollary 1) that a necessary and sufficient condition for θ to be a rational integer is to verify $\|\theta^n\| \leqq (\theta+2)^{-2}$ for $n \geqq 1$. The proof is elementary and almost trivial (an easy induction).

After, using the pigeon hole principle we give several sufficient conditions for θ to be an algebraic number. Our theorem implies a result of Pisot [3]. This result is: if $\|\lambda\theta^n\| \leqq \varepsilon$ for $n \geqq 0$ then θ is an algebraic number (in fact an algebraic integer such that all conjugates of θ have moduli at most 1). Martine Pathiaux [2] proved that the same is true if there exist q and λ, $0 \leqq q < 1/2$, $\lambda \leqq 18$, such that $\|\lambda\theta^n\| < \varepsilon_q(1+n)^{-q}$ for $n \geqq 0$ (with a small enough ε_q).

Several related results must be mentioned. Boyd [1] proved that, for example, $\|\lambda\theta^n\| \leqq 10\varepsilon$ need not even imply that θ is algebraic. Tijdeman [4] showed that for any real number $\theta > 2$ there are infinitely many λ such that $\{\lambda\theta^n\} \in [0, 1/(\theta-1)]$ for $n = 0, 1, \ldots$, (here $\{x\}$ denotes the fractional part of x).

Theorem 3 is less elementary. It can be proved without use of complex analysis, and contains older results of Pisot. We show that if the distance between $\lambda\theta^n$ and a linear recursive sequence is bounded, then θ is an algebraic integer whose conjugates have moduli $\leqq 1$, and λ and a_n are of a very special type (in the sequel, except Theorem 3, a_n is a rational integer such that $|a_n - \lambda\theta^n| \leqq 1/2$).

2. An elementary lemma.

LEMMA. *Let $L = \Sigma_{i=0}^N \alpha_i X_i$ be a linear form with integer coefficients. We suppose that*

(i) $L(a_0, \ldots, a_N) = 0$,

(ii) $\|\lambda\theta^n\| < ((\theta+1)|L|)^{-1}$ *for* $n = 0, 1, 2, \ldots$, *(where* $|L| = \Sigma|\alpha_i|$*). Then* $L(1, \theta, \ldots, \theta^N) = 0$.

Proof. Put $P(\theta) = L(1, \theta, \ldots, \theta^N)$. For all k,

$$|L(a_k, \ldots, a_{k+N}) - \lambda\theta^k P(\theta)| \leqq |L|\eta \quad (\text{where } \eta = \mathrm{Sup}_n\|\lambda\theta^n\|). \tag{1}$$

Thus, it suffices to prove that $L(a_k, \ldots, a_{k+N}) = 0$ for all k. This is true for $k = 0$. Suppose that this is still true for $k = K$ and let us prove that $L(a_{K+1}, \ldots, a_{K+1+N}) = 0$. Using (1) for $k = K$ we get $|\lambda\theta^k P(\theta)| \leqq |L|\eta$.

This inequality and (1) for $k = K+1$ give

$$|L(a_{K+1}, \ldots, a_{K+1+N})| \leqq |L|(\theta+1)\eta < 1 \quad (\text{use (ii)}).$$

The rational integer $L(a_{K+1}, \ldots, a_{K+1+N})$ is zero. The result follows by induction on k.

3. A characterization of integers.

THEOREM 1. *Let λ and θ be two positive real numbers, $\theta > 1$. We suppose that*

$$\|\lambda\theta^n\| < ((a_0 + a_1)(\theta+1))^{-1} \quad \textit{for} \quad n = 0, 1, 2 \ldots.$$

Then λ and θ are integers; $\lambda = a_0$ and $\theta = a_1/a_0$.

Proof. Using the lemma with $L = a_1X_0 - a_0X_1$, we get $\theta = a_1/a_0$ and $a_k = a_0(a_1/a_0)^k = a_0\theta^k$. Firstly, using the fact that the a_k are integers, this shows that a_0 divides a_1; θ is an integer. Secondly, the relation

$$|(\lambda - a_0)\theta^k| = |\lambda\theta^k - a_k| < 1$$

proves that $\lambda = a_0$.

The case $\lambda = 1$ gives:

COROLLARY 1. *A necessary and sufficient condition for a real number $\theta > 1$ to be an integer is that we have*

$$\|\theta^n\| < ((1+\theta)(a+1))^{-1} \quad \textit{for} \quad n \geqq 1 \quad (\textit{where} \quad |\theta - a| \leqq 1/2,\ a \in \mathbb{Z}).$$

COROLLARY 2. *A necessary and sufficient condition for a real number $\theta > 1$ to be an integer is that there exist a number k such that*

$$\|\theta^n\| \leqq ((1+\theta)^2(\theta^k + 1))^{-1} \quad \textit{for} \quad n = k, k+1, \ldots.$$

Proof. Take $\lambda = \theta^k$ in the theorem.

4. A sufficient condition to be algebraic.

THEOREM 2. *Let λ and θ be two positive real numbers, $\theta > 1$. We suppose that there exists a positive integer N such that*

$$\|\lambda\theta^n\| < ((1+\theta)(N+1)(a_0+\cdots+a_N)^{1/N})^{-1} \quad \textit{for} \quad n = 0, 1, \dots .$$

Then θ is an algebraic number of degree $\leqq N$.

Proof. We construct a linear form L such that $L(a_0, \dots, a_N) = 0$ and then apply the lemma. Consider the set of numbers $k_0a_0 + \cdots + k_Na_N$, $0 \leqq k_i \leqq H$ for $i = 0, \dots, N$, where H is a positive integer to be chosen later. These numbers belong to the interval $(0, H(a_0 + \cdots + a_N))$. Thus, two of these numbers, say, $k_0a_0 + \cdots + k_Na_N$ and $m_0a_0 + \cdots + m_Na_N$ are equal if

$$(H+1)^{N+1} > H(a_0 + \cdots + a_N) + 1.$$

Put $H = [(a_0 + \cdots + a_N)^{1/n}]$. This inequality is satisfied. So $(k_0 - m_0)x_0 + \cdots + (k_N - m_N)x_N$ is a linear form L, with integer coefficients, such that $L(a_0, \dots, a_N) = 0$ and $|L| \leqq (N+1)H$. The conclusion follows from the lemma.

If $\lambda = \theta^N$, $\theta > 1$, since $a_0 + \cdots + a_N < \theta^N(N+1)$, the theorem gives:

COROLLARY 1. *Let θ be a real number, $\theta > 1$. If there exists a positive integer N such that*

$$\|\theta^n\| < ((1+\theta)\theta^2(N-1)(N+1)^{1/N})^{-1} \quad \textit{for} \quad n \geqq N,$$

then θ is algebraic of degree $\leqq N$.

As a special case, we obtain the well-known result (see remark below):

COROLLARY 2. *If a real number $\theta > 1$ satisfies $\|\theta^n\| = o(n^{-1})$ then θ is a Pisot number.*

COROLLARY 3. *Let θ and λ be two real numbers > 1. If*

$$\|\lambda\theta^n\| \leqq (2e\theta(\theta+1)(1+\operatorname{Log}\lambda))^{-1} \quad \textit{for all} \quad n \geqq 0,$$

then θ is algebraic (of degree $\leqq \operatorname{Log}\lambda + 1$).

Take N such that $N - 1 \leqq \operatorname{Log}\lambda < N$.

5. A complement to Theorem 2. The proof of the lemma shows that (a_n) satisfies the linear relation $L(a_n, \dots, a_{n+N})$ for all $n \geqq 0$. Thus, the following result applies when the hypothesis of Theorem 2 is true.

THEOREM 3. *Let λ and θ be positive real numbers, $\theta > 1$. Suppose that there exist a sequence of integers $(a_n)_{n \geqq 0}$, such that*

$$a_n = \lambda\theta^n + O(1), \tag{2}$$

satisfying a nontrivial integer linear recurrence. Then

(i) *θ is an algebraic integer,*
(ii) *if θ' is a conjugate of θ, $\theta' \neq \theta$, then $|\theta'| \leqq 1$,*
(iii) $\lambda \in K = Q(\theta)$,
(iv) *the a_n are given by $a_n = \operatorname{Trace}_{K/Q}(\lambda\theta^n) + b_n$, where b_n is a periodic sequence of rationals.*

Proof. The proof is rather technical and will be omitted.

REMARK: The proof shows that $a_n = \lambda^n + O(1)$ implies that θ is a Pisot number.

Acknowledgement. I would like to thank the referee for several helpful remarks.

References

1. D. W. Boyd, Transcendental numbers with badly distributed powers, Proc. Amer. Math. Soc., 23 (1969) 424–427.

2. Martine Pathiaux, Répartition modulo 1 de la suite $\lambda\theta^n$, Séminaire Delange-Pisot-Poitou, 1969–1970, n°13, 6 p.

3. Ch. Pisot, Répartition (mod. 1) des puissances successives des nombres réels, Comm. Math. Helv., 19 (1946) 153–159.

4. R. Tijdeman, Note on Mahler's 3/2 — Problem, Norske Vid. Selsk., 16 (1972) 4.

UNIVERSITÉ LOUIS PASTEUR, CENTRE DE CALCUL, 7 RUE RENÉ DESCARTES, 67084 STRASBOURG, FRANCE.

RELATION OF THE CONJUGATE HARMONIC FUNCTIONS TO $f(z)$

E. V. LAITONE

In 1937 Milne-Thomson [1] pointed out the usefulness of several interesting relations between the analytic function $f(z)$ and its conjugate harmonics, $\phi(x, y)$ and $\psi(x, y)$; namely

$$f(z) = f(x + iy) = \phi(x, y) + i\psi(x, y) = \phi(z, 0) + i\psi(z, 0) \tag{1}$$

and

$$f(z) = \int [\phi_x(z, 0) - i\phi_y(z, 0)]dz = \int [\psi_y(z, 0) + i\psi_x(z, 0)]dz. \tag{2}$$

In all these relations we simply replace the real variables by $x = z$ and $y = 0$ after the required partial differentiation has been completed. For example (2) is obtained directly from the Cauchy-Riemann equations by writing

$$\begin{aligned} df/dz &= \phi_x(x, y) + i\psi_x(x, y) = \psi_y(x, y) - i\phi_y(x, y) \\ &= \phi_x(z, 0) - i\phi_y(z, 0) = \psi_y(z, 0) + i\psi_x(z, 0). \end{aligned}$$

These relations save much time and labor while also reducing the possibilities for errors since all terms containing the real variable y can be eliminated by inspection. For example, if only ϕ is given as

$$\phi(x, y) = x^6 - 15x^4y^2 + 15x^2y^4 - y^6,$$

then (2) can be immediately written as $f(z) = \int [(6z^5 + 0) - i(0)]dz = z^6$. As another example, when $\phi(x, y)$ and $\psi(x, y)$ are given by

$$\phi = \sin(x - a)\cosh(y - b); \quad \psi = \cos(x - a)\sinh(y - b), \tag{3}$$

then (1) may be written as

$$f(z) = \sin(z - a)\cosh(-b) + i\cos(z - a)\sinh(-b) = \sin(z - a - ib) = \phi + i\psi. \tag{4}$$

Despite their evident usefulness (1) and (2) have practically disappeared from current textbooks. Milne-Thomson [2] presented (1) and (2) in the 1st through 3rd (1955) editions of his text titled *Theoretical Hydrodynamics*, but then completely eliminated them in his 4th (1962) and 5th (1967) editions. He replaced (1) and (2) by a very dubious procedure that utilized the complex conjugate $\bar{z} = x - iy$. The only recent application I could find was in [3] where a short sample problem was solved by using (2) to find $f(z)$ when only $\phi(x, y)$ was known.

Because Milne-Thomson [2] had finally eliminated the much simpler procedure contained in (1) and (2), I considered the possibility that (1) and (2) may not be valid in certain cases. However, a proof for (1) and (2) can be given which delineates the validity of the method; assume that $f(z)$ is defined and analytic in a region D containing part of the real axis, say a neighborhood of $z = x_0$. Then if we define the conjugate complex function

$$\bar{f}(z) = \overline{f(\bar{z})} = \phi(x, -y) - i\psi(x, -y), \tag{5}$$

$\bar{f}$ is analytic in the region $\bar{D} = \{z : \bar{z} \in D\}$. Thus, for x real and near x_0,

$$\phi(x,0) = \frac{f(x)+\bar{f}(x)}{2}, \ \psi(x,0) = \frac{f(x)-\bar{f}(x)}{2i} \tag{6}$$

are the restrictions to a segment of the real axis of two analytic functions:

$$\phi(z,0) = \frac{f(z)+\bar{f}(z)}{2}, \ \psi(z,0) = \frac{f(z)-\bar{f}(z)}{2i}, \tag{7}$$

each defined in $D^* = D \cap \bar{D}$. Adding, we conclude that (1) is valid throughout D^*. Since the argument is local, it applies just as well to multi-valued functions, with appropriate care given to each term in (1) during analytic continuation from a neighborhood of x_0 to the full set D^*. It is easy to see from an example like

$$f(z) = \log(z - i), \ D = C - \{i\},$$

that (1) can be expected to hold only in $D^* = C - \{i, -i\}$, since

$$\psi(x, y) = \tan^{-1}\left[\frac{y-1}{x}\right] \quad \text{and} \quad \psi(z,0) = \tan^{-1}\left[\frac{-1}{z}\right] = \frac{1}{2i}\log\left[\frac{z-i}{z+i}\right].$$

To justify (2), we verify that

$$f'(z) = \phi_x(z,0) - i\phi_y(z,0) \qquad (z \in D^*), \tag{8}$$

and thus (2) holds provided the path of integration is taken in D^*. By the Cauchy-Riemann equations,

$$f'(z) = \phi_x(x, y) - i\phi_y(x, y) \qquad (z \in D) \tag{9}$$

and, if we let $F = f'$, then

$$\bar{F}(z) = \phi_x(x, -y) + i\phi_y(x, -y) \tag{10}$$

is analytic in $\bar{D}$. Thus, as for (1),

$$\phi_x(x,0) = \frac{F(x)+\bar{F}(x)}{2}, \ \phi_y(x,0) = \frac{F(x)-\bar{F}(x)}{-2i} \tag{11}$$

are restrictions to a part of the real axis of functions analytic in D^*, with

$$\phi_x(z,0) - i\phi_y(z,0) = F(z) = f'(z). \tag{12}$$

As an example of a multi-valued function, let us consider

$$\ln(z - z_0) = \ln[(x-a)^2 + (y-b)^2]^{1/2} + i\tan^{-1}[(y-b)/(x-a)]; \ (z_0 = a + ib).$$

Directly from (1) we have

$$\begin{aligned} f(z) = \phi(z,0) + i\psi(z,0) &= \ln[(z-a)^2 + b^2]^{1/2} - i\tan^{-1}[b/(z-a)] \\ &= \frac{1}{2}\ln[(z-a)^2 + b^2] + \frac{1}{2}\ln\left[\frac{(z-a)-ib}{(z-a)+ib}\right] = \ln[z - a - ib]. \end{aligned}$$

Similarly, if either ϕ or ψ is first differentiated and then substituted into (2) we obtain the correct answer, namely $f(z) = \ln(z - a - ib) + i\pi/2 + \text{CONST}$.

As a final example of a multi-valued function, let us consider

$$\phi + i\psi = (z - a)^{1/2} = (X + iy)^{1/2} = (X^2 + y^2)^{1/4}\exp(\theta/2),$$

where $\theta = \tan^{-1}(y/X)$ and $X = x - a$ so that the conjugate harmonic functions are

$$\phi(x, y) = (X^2 + y^2)^{1/4}\cos(\theta/2),\ \psi(x, y) = (X^2 + y^2)^{1/4}\sin(\theta/2).$$

Upon using (1) we immediately obtain $f(z) = [(z - a)^2]^{1/4}[\cos(0) + i\sin(0)] = (z - a)^{1/2}$. Upon differentiating the expression for ϕ we easily obtain

$$\phi_x(z, 0) = (z - a)^{-1/2}/2;\ \phi_y(z, 0) = 0$$

so (2) can be directly integrated to give $f(z) = (z - a)^{1/2}$.

In this final example, the computation time required for this procedure is now several orders of magnitude less than that required by the classical method of evaluating $\int \phi_x dy$ and/or $-\int \phi_y dx$ in order to determine $\psi(x, y)$. The solution in the form $(z - a)^{1/2}$ is even more difficult to obtain by the method presented by Milne-Thomson [**2**] in his 4th (1962) and 5th (1967) editions which may be written as

$$f(z) = 2\phi(z/2, -iz/2) + \text{CONST.} = 2i\psi(z/2, -iz/2) + \text{CONST.}$$

Milne-Thomson's method cannot be justified in general since it results in $(x^2 + y^2) = 0$, so that any $f(z)$ containing z^{-n} cannot be obtained from any combination of ϕ and/or ψ. For example, $f(z) = 1/z$ yields $\phi(x, y) = x(x^2 + y^2)^{-1}$ so that $2\phi(z/2, -iz/2) = z/0$.

In view of these examples, I believe we should retrieve (1) and (2) from their present status of near oblivion and restore them to their proper stature in texts on complex functions. At present the procedures contained in (1) and (2) give the correct answer for all textbook problems, and usually in a small fraction of the time and effort required by either the classical method or the introduction of $\bar{z}$.

Acknowledgment. The author would like to thank the referee for his comments which are presented in (5) through (12).

References

1. L. M. Milne-Thomson, On the relation of an analytic function of z to its real and imaginary parts, Math. Gazette, No. 244, 21 (1937) 228–229.

2. ———, Theoretical Hydrodynamics, Macmillan, New York, (1938, 1949, 1955, 1962, 1967), Sect. 5.32.

3. M. R. Spiegel, Complex Variables, Schaum & McGraw-Hill, New York, 1964, p. 85.

DEPARTMENT OF MECHANICAL ENGINEERING, UNIVERSITY OF CALIFORNIA, BERKELEY, CA 94720.

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics and Statistics, The University of Calgary, Calgary, Alberta, Canada, T2N 1N4. (From July 1976 to June 1977: Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, England.)

CAN THE MEASURE OF $\bigcup_1^n [a_i, b_i]$ BE COMPUTED IN LESS THAN $O(n \log n)$ STEPS?

VICTOR KLEE

Problem statement. Given n intervals $[a_1, b_1], \ldots, [a_n, b_n]$ in the real line, it is desired to find the measure of their union. How efficiently can that be done? The list of endpoints can be sorted with $O(n \log n)$ comparisons, and then $O(n)$ additions suffice to determine the measure of $\bigcup_1^n [a_i, b_i]$. However, the problem does not, *a priori*, require sorting, and it would be of interest to know whether there is a solution involving less than $O(n \log n)$ computational steps.

Reason for interest. In addition to its intrinsic appeal, the problem is of interest because it arises naturally in approximating the areas of certain subsets of the cartesian plane. Suppose that $[u_1, v_1], \ldots, [u_n, v_n]$ are intervals in the real line, f_i and g_i are continuous real functions on $[u_i, v_i]$, and

$$P_i = \{(x, y) \colon u_i \leqq x \leqq v_i, f_i(x) \leqq y \leqq g_i(x)\}.$$

Even when the areas of the individual P_i's are easily computed, it may be difficult to compute the area of the union $P = \bigcup_1^n P_i$ if the P_i's have a complicated intersection-pattern. However, when the functions are Lipschitzian the area of P is closely approximated by sums of the form

$$(*) \qquad \sum_{h=1}^{m} (\text{measure of } L_h \cap P)\,(s_h - s_{h-1}),$$

where $s_0 < s_1 < \ldots < s_m$ is an appropriate (sufficiently dense) sequence in $\bigcup_1^n [u_i, v_i]$ and L_h is the line of abscissa $(s_{h-1} + s_h)/2$. Since

$$L_h \cap P = \bigcup_{i=1}^{n} L_h \cap P_i,$$

computing (*) involves m problems of the sort mentioned above.

Solution by sorting. Suppose that each left endpoint a_i is tagged with 1 and each right endpoint b_i is tagged with -1. By means of various procedures requiring $O(n \log n)$ comparisons, the sequence of all $2n$ endpoints can be arranged in increasing order $e_1 \leqq e_2 \leqq \ldots \leqq e_{2n}$ such that if $t_1, \ldots, t_{2n}$ is the corresponding permutation of the tags, then $t_i \geqq t_{i+1}$ whenever $e_i = e_{i+1}$; that is, the sorting does not place a right endpoint before a left endpoint of the same numerical value. With EXCESS denoting the excess of the number of left endpoints over the number of right endpoints, the following program constructs the sequence $[c_1, d_1], \ldots, [c_m, d_m]$ of components of $\bigcup_1^n [a_i, b_i]$ and then computes the measure of $\bigcup_1^n [a_i, b_i]$.

```
begin
  MEASURE ← 0; m ← EXCESS ← 0;
  for i ← 1 until 2n do
```

```
    begin
      EXCESS ← EXCESS + t_i;
      if EXCESS = 1 then begin m ← m + 1; c_m ← e_i end;
      if EXCESS = 0 then d_m ← e_i
    end;
  for i ← 1 until m do MEASURE ← MEASURE + (d_i − c_i)
end
```

Problems in d-space. For d-dimensional intervals $I_1, \ldots, I_n$ in cartesian d-space, each given as the product of d real intervals, we state the following three problems and in each case ask how efficiently the problem can be solved.

PARTITION PROBLEM: *Find a sequence $H_1, \ldots, H_m$ of intervals such that $\bigcup_1^m H_k = \bigcup_1^n I_k$ and no two of the H_k's have a common interior point.*

MINIMUM PARTITION PROBLEM: *Find $H_1, \ldots, H_m$ as described such that m is minimum.*

MEASURE PROBLEM: *Find the d-measure of $\bigcup_1^n I_k$.*

When $d = 1$ the above algorithm solves all three problems in $O(n \log n)$ steps, but it is not obvious that the three problems are actually of the same computational complexity. It would be of special interest to know whether the measure problem can be solved in a number of steps that is bounded by a polynomial in d and n.

The books [1] and [2] are recommended as general references on sorting and computational complexity, while the methods of [3] can probably be applied to the partition problems.

References

1. A. V. Aho, J. E. Hopcroft and J. D. Ullman, The Design and Analysis of Computer Algorithms, Addison-Wesley, Reading, Mass., 1974.
2. D. E. Knuth, The Art of Computer Programming, Vol. 3: Sorting and Searching. Addison-Wesley, Reading, Mass., 1973.
3. H. T. Kung, F. Luccio and F. P. Preparata, On finding the maxima of a set of vectors, J. Assoc. Comp. Mach., 22 (1975) 469–476.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195.

CLASSROOM NOTES

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Laboratoire Calcul des Probabilités, Université de Paris, T.56, 4 Place Jussieu, 75–230 Paris, France.

GROUP THEORY AND THE DESIGN OF A LETTER FACING MACHINE

JOSEPH A. GALLIAN

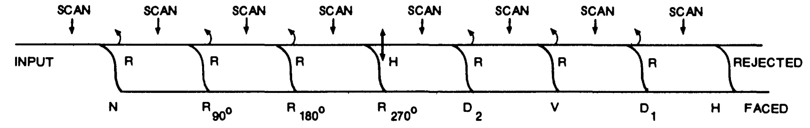
In this note we consider a number of problems relating to the optimum design of a letter facing machine for square letters. The author uses the content of this paper for the first lecture of a beginning abstract algebra course. The verifications of the solutions presented are left to the students in the course as exercises. Also, the students are asked to undertake a similar analysis for objects of other

shapes such as an equilateral triangle or a pentagon. With this approach, many elementary but important concepts of group theory are introduced in a concrete setting.

We begin by observing that any square letter moving on a conveyor can be properly faced with one of the following eight motions:

N	no change
R_{90°	rotation of 90° (counterclockwise)
R_{180°	rotation of 180°
R_{270°	rotation of 270°
H	reflection about a horizontal axis
V	reflection about vertical axis
D_1	reflection about the main diagonal
D_2	reflection about the other diagonal.

These eight motions form the dihedral group of order 8. Noting that this group is generated by $R_{90^\circ} = R$ and H we use this fact to design the machine below.



The scanners look for a stamp in the upper right hand corner of the letter. If a stamp is detected the letter is diverted to the lower track; if not, the letter continues along the upper track. The labels along the lower track indicate the net effect of the composite motions received by a letter entering the lower track at that location. A stampless letter is never diverted and ultimately ends up in a rejection bin.

For convenience, we refer to a sequence of group elements $a_1a_2\cdots a_n$ as a letter facing sequence if the set $\{a_0, a_1, a_1a_2, a_1a_2a_3, \ldots, a_1a_2\cdots a_n\}$, where a_0 is the identity, is the dihedral group of order 8. Although the above design is satisfactory from a mathematical point of view, it is not the optimum one from an engineering viewpoint. In practice, a horizontal flip is preferable to a rotation [**1**, p. 86] so the above design could be improved by utilizing five R's and two H's say. For example, the sequence RRHRRRH will also suffice to face square letters. Better yet is the sequence HRRRHRR. This sequence is more desirable than the previous one because the H's appear earlier in the sequence and as a result an "average" letter undergoes less rotation and more horizontal reflection. Specifically, on average, the sequence RRHRRRH subjects a letter to $2\frac{3}{4}$ R's and $\frac{3}{4}$ of an H while, on average, the sequence HRRRHRR subjects a letter to $2\frac{1}{4}$ R's and $1\frac{1}{4}$ H's. (This latter sequence is optimum in this regard.)

At this point it is natural to ask if it is possible to find a letter facing sequence involving R's and H's with fewer than 5 R's. The answer is "no" for sequences of length 7 and "yes" for sequences of length 8 or more. One can prove this first assertion by merely observing that a letter facing sequence of length 7 cannot possess two consecutive H's. From this it follows that any sequence of length 7 involving at least 3 H's must have a subsequence of the form HRHR or RHRH or HRRHRR. But these are each the identity element so the sequence would not generate all 8 group elements. The sequence HRHHRRHR verifies the second assertion above. In fact, among all letter facing sequences of length 8 this latter sequence is optimum with respect to having the greatest amount of horizontal flipping for the "average" letter.

We next consider the problem of finding the letter facing sequence consisting of R's and H's which is optimum with respect to involving the fewest number of R's. For example, an engineer might be happy with a letter facing sequence which consists of, say, 10 H's and only two R's. But we leave it to the reader to show that any letter facing sequence composed of R's and H's must involve at least three R's. With this in mind we now seek a letter facing sequence with 3 R's and a minimum number of H's.

Furthermore, we wish to find such a sequence with the least amount of rotation for the "average" letter. This problem has the solution HRHHRHHRH.

We summarize the above discussion in the table below. In each case we cite the letter facing sequence which has, on average, the least amount of rotation.

length of letter facing sequence	optimum design involving R and H
6 or less	impossible
7	HRRRHRR
8	HRHHRRHR
9 or more	HRHHRHHRH

Of the three motions H, R, and V, British postal engineer G. P. Copping says [**1**, pp. 86–87] that H is the most desirable and V the least. Thus, in the foregoing considerations, we have avoided using V's. Actually, the greatest expected values for H, among all letter facing sequences, are obtained by combining H's and diagonal flips. For example, the sequence $HD_1HD_1HD_1H$ has an expected value of 2 for H. This is obviously the best possible for sequences of length 7. Whether this sequence is preferable to the sequence HRRRHRR from an engineering point of view, however, is not obvious. (Copping's paper does not discuss the desirability of diagonal flips.)

Added in proof. E.G. Hills, British postal engineer, has kindly responded to an inquiry of mine about the present state of postal mechanization. According to Hills, who is the successor of the recently retired Copping, vertical reflections are now practical, even at high speed, and along with rotations and horizontal reflections, are widely used. He further adds that he knows of no way that reflections about a diagonal axis can be translated into practical equipment.

Reference

1. G. P. Copping, Automatic letter facing, British Postal Engineering, Proc. Instn. Mech. Engrs., 184 pt3H (1969–70) 85–94.

DEPARTMENT OF MATHEMATICAL SCIENCES, UNIVERSITY OF MINNESOTA-DULUTH, DULUTH, MN 55812.

ON LEBESGUE INTEGRABLE DERIVATIVES

P. L. WALKER

THEOREM. *Let f be real valued and continuous on $[a, b]$, and let D be a countable subset of $[a, b)$ such that f has a finite right hand derivative f'_+ at all points of $[a, b)\backslash D$.*

Suppose also that $f'_+ \in L^1([a, b])$. Then

$$\int_a^b f'_+(x)dx = f(b) - f(a).$$

The best known proof of this result seems to be the rather ferocious one used by Hewitt and Stromberg (in [**1**], 18.41 (d), p. 299). However, the proof via the Vitali-Carathéodory theorem, used by Rudin (in [**2**], 8.21, p. 168) in the case when $f'(x)$ is assumed to exist for *all* $x \in [a, b]$, may be extended to cover the above result by the use of the following elementary lemma.

LEMMA. *Let g be real valued and continuous on $[a, b]$, and let D be a countable subset of $[a, b)$ such that for each $x \in [a, b)\backslash D$, there exists $\delta > 0$ such that $g(y) > g(x)$ for all $y \in (x, x + \delta)$. Then g is strictly increasing on $[a, b]$.*

Proof of Lemma. Suppose firstly that there exist $c, d, a \leqq c < d \leqq b$, with $g(c) > g(d)$.

Then for each value of y with $g(c) \geqq y > g(d)$, there is a largest value of x in $[c, d)$, x_y say for which $g(x) = y$. For this value of $x = x_y$, $g(x') < g(x_y)$ for all $x' \in (x_y, d)$, and so the given condition on g fails for uncountably many points in $[c, d)$, contrary to our assumption. Hence if $a \leqq c < d \leqq b$, the $g(c) \leqq g(d)$, and g is non-decreasing on $[a, b]$.

Now let e be any point of $[c, d) \backslash D$, and let $\delta > 0$ correspond to e. Then if $x \in (e, e + \delta) \cap (e, d)$, we have $g(c) \leqq g(e) < g(x) \leqq g(d)$, and g increases strictly on $[a, b]$.

The proof of the theorem now follows closely that in [2] *loc.cit.*

Proof of the Theorem. Define $f_1(x) = f'_+(x)$ on $[a, b) \backslash D$, $= 0$ elsewhere. Then f_1 is real valued and integrable on $[a, b]$, so that given $\varepsilon > 0$, we can choose (by the Vitali-Carathéodory theorem) a lower semicontinuous function v on $[a, b]$, with $v(x) > f_1(x)$ for all $x \in [a, b]$, and $\int_a^b (v - f_1)dx < \varepsilon$.

Then if $\eta > 0$, define F_η on $[a, b]$ by

$$F_\eta(x) = \int_a^x v(t)dt - (f(x) - f(a)) + \eta(x - a).$$

Now F_η is continuous on $[a, b]$ and for all points $x \in [a, b) \backslash D$, we have

(i) there is a $\delta_1 > 0$ such that $v(y) > f_1(x)$ for all $y \in (x - \delta_1, x + \delta_1)$, and

(ii) there is a $\delta_2 > 0$ such that

$$\frac{f(y) - f(x)}{y - x} < f'_+(x) + \eta \quad \text{if} \quad y \in (x, x + \delta_2).$$

Hence if $\delta = \min(\delta_1, \delta_2) > 0$, and $y \in (x, x + \delta)$, then

$$\begin{aligned} F_\eta(y) - F_\eta(x) &= \int_x^y v(t)dt - (f(y) - f(x)) + \eta(y - x) \\ &> \int_x^y f_1(x)dt - (f'_+(x) + \eta)(y - x) + \eta(y - x) = 0. \end{aligned}$$

The Lemma now shows that F_η is strictly increasing on $[a, b]$, and in particular

$$F_\eta(b) > F_\eta(a) = 0, \quad \text{or} \quad \int_a^b v(t)dt - (f(b) - f(a)) + \eta(b - a) > 0.$$

But $\eta > 0$ is arbitrary, so $f(b) - f(a) \leqq \int_a^b v(t)dt < \int_a^b f_1(t)dt + \varepsilon$. But $\varepsilon > 0$ is also arbitrary, so $f(b) - f(a) \leqq \int_a^b f_1(t)dt = \int_a^b f'_+(t)dt$. The converse inequality now follows on applying the same reasoning to $(-f)$.

References

1. E. Hewitt and K. Stromberg, Real and Abstract Analysis, Springer Verlag, New York, 1965.
2. W. Rudin, Real and Complex Analysis, McGraw-Hill, New York, 1966.

CARTMEL COLLEGE, UNIVERSITY OF LANCASTER, BAILRIGG, LANCASTER LA1 4YL, ENGLAND.

THE SPACE OF DERIVATIVES

I. J. MADDOX

Denote by D the set of all complex valued functions f defined on $[0, 1]$ which are derivatives $f(x) = F'(x)$ of some function F, also defined on $[0, 1]$. In D we define pointwise operations, e.g. $(fg)(x) = f(x) \cdot g(x)$ for each $x \in [0, 1]$.

The main observation of this note is that D, unlike many elementary function spaces, is not an algebra.

It is a pleasure to acknowledge the very helpful comments of the referee, who suggested an improved version of my original argument and pointed out a related result in [1].

We proceed by way of a number of observations.

Observation 1. *D is a linear space (either real or complex) and contains all functions that are continuous on* $[0,1]$.

That D is linear is a trivial property of derivatives, and if f is continuous then, by the fundamental theorem of the integral calculus (for Riemann integrals),

$$\frac{d}{dx}\int_0^x f(t)dt = f(x).$$

Hence $F(x)=\int_0^x f(t)dt$ is such that $F'(x)=f(x)$.

Observation 2. *There are functions not in D, for example*

$$h(x)=x^{-1} \quad \textit{for} \quad 0<x\leqq 1;\ h(0)=0. \tag{1}$$

If $H'=h$ for some H, then H' would have to take the value 1/2 (by Darboux' theorem on derivatives), which it does not. Hence there is no such H.

Observation 3. *There are discontinuous functions in D.*

For example, $F(x)=x^2\sin(1/x^2)$, $F(0)=0$ is such that the derivative f is defined on $[0,1]$ but is not even bounded on $[0,1]$.

We now come to the main result:

Observation 4. *D is not an algebra. In fact there is a function f in D such that f^2 is not in D.*

We first note that D is closed under multiplication if and only if $f^2 \in D$ for all $f\in D$. This follows from the fact that D is a linear space and $4fg=(f+g)^2-(f-g)^2$.

Now for $x>0$ define the functions:

$$G(x)=x^2\exp(i/x^2),\ k(x)=2x\exp(i/x^2),$$

$$f(x)=x^{-1}\exp(i/x^2),\ g(x)=x\exp(-2i/x^2).$$

Also, define $G(0)=k(0)=f(0)=g(0)=0$.

A computation yields $G'=k-2if$ on $[0,1]$. Now $G'\in D$, and since k is continuous, we have $k\in D$. Thus, since D is a linear space, it follows that $f\in D$.

Also, since g is continuous, we have $g\in D$. Hence, if we suppose that D is an algebra, then we would have $gf^2\in D$. But $gf^2=h$, where h is defined by (1) above, and this contradicts the fact that h is not in D. Thus D cannot be an algebra.

A similar argument may be given that uses only real valued functions. One shows that $F\in D$, where we define $F(x)=x^{-1/2}\cos(1/x^2)$ if $x>0$, and $F(0)=0$. Then, assuming $F^2\in D$, we find that h, defined by (1), is in D, which is false.

Reference

1. J. M. Foran, On the product of derivatives, Fund. Math., 80 (1973) 293–294, MR 48 # 4218.

Department of Pure Mathematics, Queen's University, Belfast BT7 1NN, N. Ireland.

EDITED BY PAUL T. MIELKE AND SHIRLEY HILL

Material for this Department should be sent to Paul T. Mielke, Department of Mathematics, Wabash College, Crawfordsville, IN 47933, or to Shirley Hill, Department of Mathematics, University of Missouri, Kansas City, MO 64110.

A LEARNING CENTER BASED SYSTEM OF INSTRUCTION

GARY MUSSER AND LINDA THOMPSON

1. Introduction. For a variety of reasons the level and quality of the mathematical preparation of beginning college students have been declining during the past several years. Hence, there has been a corresponding increase in the number of students enrolling in courses preceding the calculus (intermediate algebra, college algebra, trigonometry, and elementary functions) at Oregon State University. In an effort to shift more of the instructional responsibility onto the students in these courses, a modest Mathematics Learning Center (MLC) was initiated in 1973 and an expanded facility was opened in the winter term of 1975. Although the MLC offers a variety of services to both faculty and students, this article is devoted to a description of our evolving learning center based system of instruction and its innovative testing component.

2. History. During the learning center's initial year of operation one section (30–50 students) of college algebra was offered each term via an independent study mode of instruction. These experimental sections were selected by students in preference to conventional sections mainly due to the self-paced aspect of the course. The course content was organized into eight modules and students were provided with the objectives and assignments for each module. The students studied from a conventional textbook; however, a variety of support materials including audiocasette tapes, workbooks, and reference texts were available in the MLC. Upon completion of a module the students took a multiple choice mastery test which was graded immediately by the course instructor (a graduate student in mathematics education) who was in the MLC during certain scheduled hours. When a student had completed all eight modules he was given a comprehensive final exam. Finally, although the students were self-paced, a time schedule was provided to minimize procrastination and all students were required to finish the course by the end of the term.

Although this instructional model has many features of the Personalized System of Instruction model (PSI), or Keller Plan [1], it also differed in several ways. First, all students were required to complete the entire course since the material covered would be of use to them in their future coursework. Secondly, the ten-item mastery quizzes had a multiple choice format and were graded according to the following scale: 70% – C, 80% – B, and 90% or 100% – A; no partial credit was given. A student had to score at least 70% on a quiz to be allowed to take a quiz on the next module, thus a level of mastery was imposed. Several parallel forms of each quiz were available for retesting. The multiple choice format was used since only the course instructor was available to grade quizzes and provide feedback. Finally, as a quality control check, students who needed the entire term to finish the course (about 70% of the students) took a common mass final exam with all students from the conventional sections. Based on a comparison of these final exam grades, students studying under this method seemed to do as well as or better than students in the conventional sections. However, it seemed that it was the better prepared students who elected this mode of instruction.

3. Learning center based system of instruction. After this model had been in use on a pilot basis for one year the MLC moved to its larger, remodeled facility. This facility was open 48 hours per week and there were usually 3 consultants (faculty, graduate students, or undergraduate workstudy students) available to assist students. In addition there was a full-time clerk to circulate audio-tutorial

materials, administer and grade quizzes, and keep records. Because these qualified personnel were available throughout the day for both assistance and testing, the instructional model was modified to allow more flexibility by permitting the students to take quizzes anytime during the day. As a result, many students would rarely see their instructor and instead would use the MLC mainly as a testing center. Even though they could see a consultant for instruction concerning missed questions, many students did not avail themselves of this opportunity.

Since this modified instructional model had the potential of serving more students per faculty member than did a conventional classroom format (thus permitting a more favorable student-teacher ratio in our university level courses), the model was desirable from the standpoint of economics. However, two desirable components of PSI, namely the proctorial aspect (10 students per proctor) and the open-ended objective testing feature, were impossible to administer with the economy of our model.

Our model required a testing format that could be administered and scored quickly and easily by people not necessarily expert in mathematics. Thus, in an effort to preserve the efficiency of our model while incorporating the positive ingredients of the testing and feedback aspects of the PSI model, a new multiple choice testing procedure was implemented and then evaluated.

4. Answer-until-correct testing. The idea of testing with immediate feedback on the correctness of the answers can be traced back to S. L. Pressey [2] in the 1920's, but, as is common, this concept was ahead of the technology. Now, however, there is a convenient process called a latent image developer which permits immediate feedback in testing. After test items have been typed on a spirit (ditto) master, a chemically treated latent image transfer sheet (available through the A. B. Dick Company) is used to insert invisible marks on the master before multiple copies are run off. A specially prepared felt tip marker is then used by the student to bring out the latent image on the copies.

In the independent study sections, the following quiz format was used:

1. $5! =$

——A. 15 ——B. 120 ——C. 20 ——D. 60.

Latent images were placed on each blank so that a " + " would appear if the correct answer blank was marked and a "0" if an incorrect one was marked. A student would continue marking until the correct answer was identified by the +. Then, partial credit was assigned depending on the number of attempts needed to determine the correct answer: 10 points for one attempt, 5 points for two, 2 points for three, and 0 points otherwise.

Because student reaction to the answer-until-correct (a-u-c) format was favorable and its self-scoring attribute made it efficient to use in the MLC, a research study was carried out in the Spring 1975 term to obtain information on its effect on achievement, its level of anxiety production and its reliability, all relative to the standard multiple choice (s-m-c) format.

To see if the a-u-c format was superior to the s-m-c format college algebra students in an independent study section were assigned randomly to two groups. One group used the a-u-c format on all quizzes up to the midterm exam and the other used the s-m-c. An open-ended midterm exam was administered to the entire class and was graded on a basis of 200 points by a jury of three instructors not associated with the experiment. The results are in Table 1.

TABLE 1: A-U-C and S-M-C Midterm Exam Achievement

Group	Mean	Standard Deviation	Number of Students
a-u-c	137	32.4	15
s-m-c	123.5	45.6	21

A t-test showed that this difference in achievement favoring the a-u-c format is significant at the .14 level.

To determine the anxiety level of a-u-c relative to s-m-c the same two groups were asked to respond to the five point Likert-type scale shown in Figure 1.

We are interested in determining how the format of your quizzes in this course affect your nervousness or anxiety towards them. Please mark on the scale below your general anxiety level while you are taking Math 101S quizzes.

1	2	3	4	5
Very Calm	Relatively Calm		Anxious	Very Anxious

FIG. 1.

The results from this form are in Table 2.

TABLE 2: A-U-C and S-M-C Anxiety

Group	Mean	Standard Deviation	Number of Students
a-u-c	2.53	1.11	15
s-m-c	3.00	1.25	21

A t-test shows that this difference in anxiety favoring the a-u-c format is significant at the .27 level.

To test the reliability of the a-u-c format relative to the s-m-c format the students in the intermediate algebra independent study section were randomly assigned to two groups. Each group was administered a midterm exam, half open-ended and half multiple-choice, where the two halves were judged to be comparable in both content and difficulty. To account for possible teaching interaction due to the testing procedure, the two groups were divided and tested as follows:

Group 1a: a-u-c first, open-ended second
Group 1b: open-ended first, a-u-c second
Group 2a: s-m-c first, open-ended second
Group 2b: open-ended first, s-m-c second

These results are summarized in Table 3.

TABLE 3: Correlations

Group	Number	Correlation between two parts of the exam
1	19	.77
2	24	.76

Although there is no significant difference between these correlations at the .01 level, the relatively high correlations show that both a-u-c and s-m-c are reliable measures of achievement when compared with an open-ended objective test.

In summary, this study shows that a-u-c and s-m-c are comparable and suggests that a-u-c may even be a better format for testing. Thus, the use of a-u-c format should be continued and even expanded, especially because of its efficiency and popularity with students.

5. Future plans. Even though the a-u-c format improves the feedback aspect associated with testing, students still need additional encouragement to obtain assistance when weaknesses have been identified by a quiz. Therefore, beginning with the 1975 fall term when we expect to have over 200 students in the program, a student making more than two marks on any quiz problem will be encouraged to visit with a consultant regarding the student's misconception associated with that problem. Thus, this modified model better approximates the proctorial aspect of the PSI model while maintaining the economy of the learning center based instructional model.

Thus far, our independent study sections have been restricted to intermediate algebra, college

algebra, and trigonometry. However, during the summer of 1975 materials were prepared under a grant from the Oregon Educational Coordinating Council to permit an extension of the mode of instruction to our elementary functions course for calculus preparation. Nine modules were prepared to serve as a content core for the course. In addition, a cassette-tape and booklet package was prepared for each module to assist students in developing their problem-solving skills. Finally, some brief videocassettes (about 15 minutes) will be prepared to provide a preview or review for each module. It is expected that these videocassettes will effectively summarize the content central to the module. It is hoped that this instructional package will accommodate a wide variance in student backgrounds and abilities.

6. Evaluation. Since this learning center based instructional model is still evolving the entire model has not been thoroughly evaluated and there are several questions that can be asked. Do students electing this mode of instruction achieve as much on a common final exam as their counterparts in conventional sections? Do they retain it as well and apply it as well as students in conventional sections in the long term? Is there a certain kind of student that achieves more under this system? What is the student attitude towards this mode of instruction? Do students take more or less mathematics when this option is available? Are attitudes towards mathematics affected? Is this method more cost efficient? What is the maximum number of students that can be served effectively by one instructor? Are more responsible study habits nurtured by this method?

Some of these questions will be answered by an internal evaluation currently being conducted by one of our colleagues. It is hoped that frequent evaluation and the implementation of modifications suggested by this feedback will produce a "final" learning center based instructional model which will be at least as effective as our conventional classes, and moreover, one which is more efficient in terms of both student and faculty time.

References

1. F. S. Keller and J. G. Sherman, The Keller Plan Handbook, Benjamin, Menlo Park, 1974.

2. S. L. Pressey, A simple apparatus which gives tests and scores and teaches, School and Society 23: 373–76, March 20, 1926.

3. L. G. Thompson, A study of the effect of an answer-until-correct multiple-choice procedure on mathematics achievement, unpublished doctoral dissertation, Oregon State University, 1975.

DEPARTMENT OF MATHEMATICS, OREGON STATE UNIVERSITY, CORVALLIS, OR 97331.

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before July 31, 1977.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

E 2647. *Proposed by Daniel Gallin, University of San Francisco*

Let Γ_1, Γ_2 be two continuous maps of the unit segment $I = \{x \mid 0 \leqq x \leqq 1\}$ into the unit square I^2. Suppose that $\Gamma_1(0) = (0,0)$, $\Gamma_1(1) = (1,1)$, $\Gamma_2(0) = (0,1)$, $\Gamma_2(1) = (1,0)$. Prove by elementary means (e.g., without using the Jordan Curve Theorem) that the two curves Γ_1 and Γ_2 meet.

E 2648. *Proposed (part (i)) by R. P. Nederpelt, Eindhoven University of Technology, Netherlands, and (part (ii)) by R. B. Eggleton, Northern Illinois University, and John H. Loxton, University of New South Wales, Australia*

(i) Show that there is no infinite sequence of prime numbers $p_1, p_2, \ldots$ such that $p_{k+1} = 2p_k \pm 1$ for all k.

(ii)* Find a longest finite sequence of primes $p_1, p_2, \ldots, p_n$ such that $p_{k+1} = 2p_k + 1$ for $1 \leqq k \leqq n-1$.

E 2649. *Proposed by A. Oppenheim, University of Benin, Nigeria*

Let a, b, c and α, β, γ be the sides and the corresponding opposite angles of a non-obtuse triangle. Show that

$$3(a+b+c) \leqq \pi\left(\frac{a}{\alpha}+\frac{b}{\beta}+\frac{c}{\gamma}\right), \tag{1}$$

$$3(a^2+b^2+c^2) \geqq \pi\left(\frac{a^2}{\alpha}+\frac{b^2}{\beta}+\frac{c^2}{\gamma}\right). \tag{2}$$

E 2650. *Proposed by M. J. Pelling, University of Benin, Nigeria*

Find the Galois group of the equation $x^9 + x^3 + 1 = 0$ over the rationals.

E 2651. *Proposed by Paul Erdös*

A finite number of pennies are placed flat on the plane so that no two overlap and no three touch each other. Prove that these pennies can be painted with at most three colors so that touching pennies bear different colors. (This is a variant of Problem E 2527 [1976, 485].)

E 2652. *Proposed by Jeffrey L. Rackusin, California State University at Northridge*

Let $A = (a_{ij})$ be a row-stochastic $n \times n$ matrix. Show that

$$\sum_{\sigma \in S_n} \prod_{i=1}^{n} \left(a_{i,\sigma(i)} \Big/ \sum_{j=1}^{n} a_{i,\sigma(j)} \right) = 1,$$

where S_n is the symmetric group.

SOLUTIONS OF ELEMENTARY PROBLEMS

A Bernoulli Differential Equation

E 2568 [1975, 1010]. *Proposed by Stroughton Bell, University of New Mexico*

Show that the Bernoulli equation

$$y' + y^2 + xy = 0$$

has exactly two solutions on the entire line for which y'' is nowhere zero.

Solution by Samuel Ray Glidewell and Joseph Fort Grimland, Jr., Georgia Institute of Technology, with simplifications due to several other solvers. Suppose y is a solution to the problem. Then $y'' = y(2y^2 + 3xy + x^2 - 1)$. Hence y never vanishes and $(1/y)' = x(1/y) + 1$, so

$$\frac{1}{y} = e^{x^2/2}\left(C + \int_0^x e^{-t^2/2} dt\right) \quad \text{with} \quad |C| \geqq \sqrt{\pi/2}.$$

We cannot have $|C| > \sqrt{\pi/2}$ since in this case $\lim_{|x| \to \infty} y = 0$, so the graph of y would intersect the hyperbola $2y^2 + 3xy + x^2 = 1$ contradicting the assumption that y'' never vanishes. Thus the problem has at most two solutions.

If y_1 and y_2 are solutions to the Bernoulli equation corresponding to $C = \sqrt{\pi/2}$ and $C = -\sqrt{\pi/2}$ respectively then $y_1(x) = -y_2(-x)$. It thus suffices to show that y'' never vanishes in case

$$\frac{1}{y} = e^{x^2/2}\left(\sqrt{\pi/2} + \int_0^x e^{-t^2/2} dt\right) = e^{x^2/2} \int_{-\infty}^x e^{-t^2/2} dt.$$

This will be the case if and only if the graph of y never intersects the hyperbola $2y^2 + 3xy + x^2 = 1$ and this will certainly be true if we can show that

$$\text{(1)} \qquad \frac{e^{-x^2/2}}{\int_{-\infty}^x e^{-t^2/2} dt} > \frac{-3x + \sqrt{x^2 + 8}}{4}$$

for all real x. This inequality is clearly true if $x \geqq 1$ since then $-3x + \sqrt{x^2 + 8} \leqq 0$. Let

$$F(x) = \frac{4e^{-x^2/2}}{-3x + \sqrt{x^2 + 8}} - \int_{-\infty}^x e^{-t^2/2} dt \quad \text{for} \quad x < 1.$$

Then $F'(x) = 2e^{-x^2/2}\, g(x)(x^2+8)^{-1/2}(-3x+\sqrt{x^2+8})^{-2}$ where $g(x) = (x^2+2)\sqrt{x^2+8} + x(x^2+6)$. It is easy to verify that $g(x) > 0$ for all real x. Therefore $F'(x) > 0$ for $x < 1$. But $\lim_{x\to-\infty} F(x) = 0$ and so $F(x) > 0$ for $x < 1$. Hence (1) holds also when $x < 1$ and we are done.

Also solved by A. M. J. Davis (England), L. Kuipers (Switzerland), David Neu, and the proposer. That the problem has at most two solutions was shown by Gary Gunderson, O. P. Lossers (Netherlands), and the St. Olaf Problem Group. D. K. Cohoon showed that the analogous problem for the equation $y' + x^{2n-1}y + y^2 = 0$ has at most two solutions.

Stack of Pancakes

E 2569* [1975, 1010]. *Proposed by Harry Dweighter, The City College of the City University of New York*

The chef in our place is sloppy, and when he prepares a stack of pancakes they come out all different sizes. Therefore, when I deliver them to a customer, on the way to the table I rearrange them (so that the smallest winds up on top, and so on, down to the largest on the bottom) by grabbing several from the top and flipping them over, repeating this (varying the number I flip) as many times as necessary. If there are n pancakes, what is the maximum number of flips (as a function of n) that I shall ever have to use to rearrange them?

Comments by Michael R. Garey, David S. Johnson and Shen Lin, Bell Labs, Murray Hill, New Jersey. Let $f(n)$ be the required number. By hand and computer search we have found the following values:

n	1	2	3	4	5	6	7
$f(n)$	0	1	3	4	5	7	8

We can also show that $n+1 \leqq f(n) \leqq 2n-6$ for $n \geqq 7$.

Further contributions are solicited.

Lattice Points and Least Common Multiple

E 2570 [1976, 53]. *Proposed by J. G. Sunday, University of Guelph, Ontario, Canada*

Let $(m_1, n_1), \ldots, (m_k, n_k)$ be distinct lattice points with $n_i \geqq 2m_i > 0$ for each i, and suppose that no two of them lie on any line through the origin. Show that the least common multiple of $n_1, \ldots, n_k$ is not less than $2k$. When can equality occur?

Solution by Mark Kleiman, Student, Stuyvesant High School, Staten Island, N.Y. Let a be the least common multiple of $n_1, \ldots, n_k$ and let $x_i = a/n_i$. Then $0 < m_i x_i \leqq a/2$. The hypothesis of the problem implies that $m_1x_1, \ldots, m_kx_k$ are k distinct integers. Therefore $a \geqq 2k$. Equality holds if and only if a is even and for every integer b such that $a \geqq 2b > 0$ there exists a point (m_i, n_i) lying on the segment joining (a, b) to the origin.

Also solved by David Bienenfeld (Israel), J. M. Brown & David Voss, Peter de Buda, A. J. Douglas (England), Donald Fuller, Paul Garlick, Richard Gibbs, Landy Godbold, M. G. Greening (Australia), Lael Kinch, Jordan Levy, L. E. Mattics, Ram Murty & Kumar Murty (Canada), Mark Passolt, and the proposer.

Perfect-plus-two Numbers

E 2571 [1976, 53]. *Proposed by Sidney Kravitz, Dover, New Jersey*

It is well known that if $2^p - 1$ is prime, then $n = 2^{p-1}(2^p - 1)$ is a perfect number (i.e., $\sigma(n) = 2n$

where $\sigma(n)$ is the sum of the divisors of n) and that every even perfect number is of this form. A number n is perfect-plus-one (pp1) if $\sigma(n) = 2n - 1$; see Problem E 1445* [1960, 1028; 1975, 73] and the following references; [**1**], [**2**], listed at the end of this solution. It is known that if $n = 2^k$ then n is pp1, but it is not known if there are any other pp1 numbers.

Discuss the situation for pp2 numbers, i.e., numbers n for which $\sigma(n) = 2n - 2$.

Discussion by K. Inkeri, University of Turku, Finland. The numbers $2^{k-1}(2^k - 1)$, with $2^k - 1$ prime, are pp0 (= perfect). The numbers 2^k are pp1. The numbers $2^{k-1}(2^k + 1)$, with $2^k + 1$ prime, are pp2. No other pp0, pp1 or pp2 numbers are known.

THEOREM 1. *All the pp2 numbers, which have at most two distinct prime factors, are given by $n = 2^{k-1}(2^k + 1)$ where $2^k + 1$ is prime $(k \geqq 1)$.*

Proof. For prime p and $h \geqq 1$ we have

$$\frac{p+1}{p} \leqq \frac{\sigma(p^h)}{p^h} = \frac{p - p^{-h}}{p-1} < \frac{p}{p-1}. \tag{1}$$

Let n be a pp2 number having at most two distinct prime factors. Assume first that $n = 2^a p^b$ where p is an odd prime. If $b = 0$ then $n = 2^a$ is a pp1 number. Therefore $b \geqq 1$. Using (1) we obtain

$$(2^{a+1} - 1)\frac{p+1}{p} \leqq \frac{\sigma(n)}{p^b} = 2^{a+1} - \frac{2}{p^b} < (2^{a+1} - 1)\frac{p}{p-1}. \tag{2}$$

From $(2^{a+1} - 1)(1 + 1/p) < 2^{a+1}$, it follows that $p > 2^{a+1} - 1$. Since p is an odd prime we must have $p \geqq 2^{a+1} + 1$. The second inequality in (2) gives $p < 2^{a+1} + 2(p-1)p^{-b}$. Hence we must have $2(p-1) > p^b$, and consequently $b = 1$ and $p = 2^{a+1} + 1$.

It remains to show that $n = p^a q^b$ $(a \geqq 1,\ b \geqq 1)$, where p and q are two distinct odd primes, cannot be pp2. By (1) we have

$$2 - \frac{2}{n} = \frac{\sigma(n)}{n} = \frac{\sigma(p^a)}{p^a}\cdot\frac{\sigma(q^b)}{q^b} < \frac{p}{p-1}\cdot\frac{q}{q-1}$$

which can be written as

$$\frac{1}{p-1} + \frac{1}{q-1} + \frac{1}{(p-1)(q-1)} + \frac{2}{n} > 1. \tag{3}$$

We may assume that $p < q$. Then (3) implies that $p = 3$. Now (3) gives

$$\frac{3}{q-1} + \frac{4}{n} > 1$$

and it follows that $q = 5$ and $a = b = 1$. Thus $n = 15$, but it is not pp2.

The following readers show that the numbers in the theorem are pp2: Fred Buckley, John DeCarlo, Thomas Elsner, Richard Gibbs, Lael Kinch, L. Kuipers (Switzerland), R. Laumen (Belgium), and S. C. Locke (Canada).

Inkeri proves also that if an odd number is pp2, then it must be divisible by at least 4 distinct primes. Kinch points out that if n odd is pp2 and $3 \nmid n$ then n must be divisible by at least 7 distinct primes. DeCarlo asks if there are any pp3 numbers.

References

1. R. P. Jerrard and Nicholas Temperley, Almost perfect numbers, Math. Mag. 46 (1973), 84–87.
2. J. T. Cross, A note on almost perfect numbers, Math. Mag. 47 (1974), 230–231.

Can a Derivative be Differentiable at a Limit Point of its Discontinuities?

E 2572 [1976, 53]. *Proposed by C. D. H. Cooper, Macquarie University, North Ryde, Australia*

Prove, or give a counterexample: If $f: R \to R$ is differentiable everywhere and f' is differentiable at some point a, then f' is continuous in some neighborhood of a.

Solution by Thomas Sellke, Wabash College, and Emanuel Schorsch, Indiana University (independently). The assertion is false. The following is a counterexample. Let $f(x) = 0$ if $x \leqq 0$, $x \geqq 1$, or $x = 1/n$ for a positive integer n; and let

$$f(x) = \left(x - \frac{1}{n+1}\right)^2 \left(x - \frac{1}{n}\right)^2 \sin \frac{1}{x - (1/(n+1))}$$

for $x \in ((n+1)^{-1}, n^{-1})$. It is easily shown that f is everywhere differentiable, that f' is discontinuous at every point $1/n$, and that $f''(0)$ exists and is equal to zero.

Also solved by Irl Bivens, John Bryant, Gerd Fricke, William Gorman III, William Habakkuk, G. A. Heuer, L. E. Mattics, Alan Shuchat, Arthur Solomon, R. H. Sorgenfrey, Jón Stefánsson (Iceland), Richard Stevens, R. M. Warten, William Weimer, and the proposer. Partial solutions by D. M. Bloom, and Guillermo Hansen (Argentina).

Editor's comment. Fricke gives an example of a function f which is differentiable everywhere, and which has the property that f' is differentiable on a dense subset of R and discontinuous on another dense subset of R. Stefánnson notes that the set of discontinuity points of the derivative, being a pointwise limit of continuous functions, always is a meagre set and that the set of continuity points is everywhere dense. [R. P. Boas, *A Primer of Real Functions*, The Mathematical Association of America, 1972, pp. 103–106.] The set of discontinuity points of the derivative need not be of measure zero, according to an example of Volterra. [B. Sz.-Nagy, *Introduction to Real Functions and Orthogonal Expansions*. Oxford University Press, 1965, p. 154.]

Inequalities for Symmetric Functions

E 2573 [1976, 54]. *Proposed by Murray S. Klamkin, University of Waterloo, Ontario*

If n positive real numbers vary such that the sum of their reciprocals is fixed and equal to A, find the maximum value of the sum of the reciprocals of the $\binom{n}{j}$ sums of the n numbers taken j at a time.

I. *Solution by Russell Lyons, undergraduate, Case Western Reserve University.* Let $a_1, \ldots, a_n$ be the n numbers. By the theorem of the arithmetic and harmonic means,

$$\frac{1}{j}\sum_{i=1}^{j} a_i \geqq \frac{j}{\sum_{i=1}^{j} 1/a_i} \quad \text{or} \quad \frac{1}{a_1 + \cdots + a_j} \leqq \frac{1}{j^2}\sum_{i=1}^{j}\frac{1}{a_i},$$

with equality if and only if $a_1 = a_2 = \cdots = a_j$. A similar inequality holds, of course, for any j of the n numbers. Summing all such $\binom{n}{j}$ inequalities, we have

$$(1) \qquad \sum_{\text{sym}} \frac{1}{a_1 + \cdots + a_j} \leqq \frac{1}{j^2}\binom{n-1}{j-1}\sum_{i=1}^{n}\frac{1}{a_i} = \frac{1}{j^2}\binom{n-1}{j-1}A,$$

with equality if and only if $a_1 = a_2 = \cdots = a_n$.

II. *Solution by A. McD. Mercer, University of Guelph, Ontario.* If $x_1, \ldots, x_n$ are non-negative then

$$jx_{k_1}x_{k_2}\cdots x_{k_j} \leqq x_{k_1}^j + x_{k_2}^j + \cdots + x_{k_j}^j.$$

Summing these inequalities for $1 \leqq k_1 < k_2 < \cdots < k_j \leqq n$ we get

$$j\sum_{\text{sym}} x_1x_2\cdots x_j \leqq \binom{n-1}{j-1}\sum_{k=1}^{n} x_k^j.$$

Now put $x_k = \exp(-a_k t)$, where $a_k > 0$, and integrate with respect to t over $(0, +\infty)$ to get (1).

III. *Solution and a generalization by John Williams, NSW, Australia (revised by the editor).* We claim that if $x_1, \ldots, x_n$ are positive real numbers such that $\Sigma_1^n 1/x_i = A$ and $a \geqq 0$ then

$$\max \sum_{\text{sym}} \frac{1}{x_1 + \cdots + x_j + a} = \binom{n}{j} \frac{A}{nj + aA}. \tag{2}$$

Proof. The set of all n-tuples $(x_1, \ldots, x_n)$ satisfying the above conditions is a compact subset K of $\mathbf{R}^n$. Let $(a_1, \ldots, a_n) \in K$ be a point where the function (defined on K)

$$F(x_1, \ldots, x_n) = \sum_{\text{sym}} \frac{1}{x_1 + \cdots + x_j + a}$$

attains its maximum. It is easy to verify that if $x_1 \neq x_2$ then

$$F(x_1, x_2, \ldots, x_n) < F(y, y, x_3, \ldots, x_n)$$

where $y = (x_1 + x_2)/2$. This inequality implies that $a_1 = a_2$. Similarly, $a_1 = a_2 = \cdots = a_n = n/A$. Thus (2) follows by evaluating F at this point.

IV. *Solution and a generalization by the proposer.* Let $a_1, \ldots, a_n$ be positive and define

$$S_r = \sum_{\text{sym}} \frac{1}{x_1 + \cdots + x_r} \quad \text{and} \quad T_r = \frac{r^2 S_r}{\binom{n-1}{r-1}}.$$

We shall prove that $T_1 \geqq T_2 \geqq \cdots \geqq T_n$.

Proof. The inequality

$$T_{n-1} = (n-1)S_{n-1} \geqq n^2 S_n = T_n \tag{3}$$

follows easily from the Cauchy-Schwarz inequality because

$$\frac{n-1}{S_n} = (n-1)(x_1 + \cdots + x_n) = \sum_{\text{sym}} (x_1 + \cdots + x_{n-1}).$$

Using (3), we have $(r = 2, 3, \ldots, n-1)$

$$(r-1) \sum_{\text{sym}} \frac{1}{y_1 + \cdots + y_{r-1}} \geqq \frac{r^2}{y_1 + \cdots + y_r}$$

for any positive $y_1, y_2, \ldots, y_r$. Replacing $(y_1, \ldots, y_r)$ by $(x_{k_1}, \ldots, x_{k_r})$ and summing over all r-tuples $(k_1, \ldots, k_r)$ such that $1 \leqq k_1 < \cdots < k_r \leqq n$ we obtain $(r-1)(n-r+1)S_{r-1} \geqq r^2 S_r$ i.e., $T_{r-1} \geqq T_r$.

Note that the proposed inequality is identical with $T_1 \geqq T_j$.

Also solved by Peter de Buda, Edward Dixon, L. Kuipers (Switzerland), Joel Levy, L. E. Mattics, and Problem Solving Group Bern (Switzerland).

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Hill Center, Rutgers University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate signed sheets and should be mailed before July 31, 1977.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

6144. *Proposed by Carl Pomerance, University of Georgia*

If n is a natural number, denote by $A(n)$ the arithmetic mean of the divisors of n. (See O. Ore, *On*

the averages of the divisors of a number, this MONTHLY 55 (1948), 615–619.)

(a) Prove that the asymptotic density of the set of n, for which $A(n)$ is an integer, is 1.

(b) Show that for any N there is an integer m such that $A(n)=m$ has at least N solutions.

(c)* If it exists, find the asymptotic density of the set of integers m for which $A(n)=m$ has a solution.

6145. *Proposed by Michael Barr, Eidgenössische Technische Hochschule, Zürich, Switzerland*

Let $\mathbb{N}=\{0,1,2,\ldots\}$ be the natural numbers, $\mathbb{C}^*$ the nonzero complex numbers. Suppose $\rho:\mathbb{N}\times\mathbb{N}\to\mathbb{C}^*$ is a "kernel function" with the property that the convolution product defined on functions $\mathbb{N}\to\mathbb{C}^*$ by the formula

$$(f *_\rho g)(n) = \sum_{i+j=n} \rho(i,j)f(i)g(j)$$

is associative. Show that there is a function $\sigma:\mathbb{N}\to\mathbb{C}^*$ such that $f *_\rho g = \sigma^{-1}(\sigma f *_\sigma g)$, where an unadorned $*$ denotes the usual convolution with respect to the kernel which is identically 1. Note that this implies that $f *_\rho g = g *_\rho f$ and ultimately that ρ is symmetric.

6146. *Proposed by Edward J. Wegman, University of North Carolina and Anton Glaser, Pennsylvania State University, Abington Campus*

The year 1623 marked not only the publication of Shakespeare's *First Folio*, but also of Sir Francis Bacon's *De Augmentis Scientiarum*, in which Bacon proudly explained his "biliteral" cipher for writing "*omnia per omnia*" or anything by anything. Bacon assigned the 24 letters of the alphabet (j and u were absent) to the first 24 5-bit strings from 00000 to 10111. (Actually he used the 5-letter words AAAAA to BABBB, but this is a non-essential difference.) The word Bacon would appear as

00001 00000 00010 01101 01100

and this in turn could be hidden in a covertext of at least 25 letters, such as

00	00	10	000	00	00	1001	10	101	100
↓ ↓	↓ ↓	↓ ↓	↓ ↓ ↓	↓ ↓	↓ ↓	↓ ↓ ↓ ↓	↓ ↓	↓ ↓ ↓	↓ ↓ ↓
To	be,	*o*r	not	to	be,	*t*ha*t*	*i*s	*t*h*e*	*q*uestion.

Here, "0" was replaced by one typestyle (in this case Roman) and "1" by another (in this case *Italic*). Thus, the 4,500,000 letters of the *First Folio* may be interpreted as a string of 4,500,000 binary digits.

What is the probability that the message "Bacon wrote this" appears in the *First Folio* "by accident"

(a) if the probability of a letter's being Roman is $\frac{1}{2}$?

(b) if the probability of the letter's being Italic is 1/10?

6147. *Proposed by Richard Johnsonbaugh, Chicago State University*

Can a normal, separable space possess a closed, uncountable, discrete subspace?

6148. *Proposed by Charles Small, Queen's University, Kingston, Ontario, Canada*

Let $s(n)$ denote the smallest r such that -1 is a sum of r squares mod n. (Thus $s(n)$ is the *Stufe*, or "level", of the ring $Z/(n)$.) Show that $s(n)$ is computed as follows:

$s(n)=1$ if $4 \nmid n$ and $p \nmid n$ for all primes $p \equiv 3 \pmod 4$;

$s(n)=2$ if $4 \nmid n$ and $p \mid n$ for some prime $p \equiv 3 \pmod 4$;

$s(n)=3$ if $4 \mid n$ but $8 \nmid n$;

$s(n)=4$ if $8 \mid n$.

6149. *Proposed by Gérard Letac, Université Paul-Sabatier, Toulouse, France*

A bug runs along the edges of a regular dodecahedron with constant speed: one edge per unit of time. At time 0 the bug is on some vertex A; at time n (n an integer) it chooses randomly one of the three possible edges. If p_n is the probability that the bug is on A at time n, then it is trivial to compute $p_0 = 1$, $p_1 = 0$, $p_2 = \frac{1}{3}$, $p_3 = 0$, $p_4 = 5/27, \ldots$. Determine the generating function $\Sigma_{n=0}^{\infty} p_n s^n$ of the sequence $\{p_n\}$.

SOLUTIONS OF ADVANCED PROBLEMS

Power Series for which $f'(r) > 0$

6038 [1975, 671]. *Proposed by Oto Strauch, Bratislava, Czechoslovakia*

Let $f(x) = \Sigma a_j x^j$ and $s_n(x) = \Sigma_{i \leq n} a_j x^j$. Let $r \neq 0$ be an interior point of the interval of convergence of the power series $\Sigma a_j x^j$. Prove that if $s_n(r) < f(r)$ for every $n = 0, 1, 2, \ldots$, then the derivative $f'(r) \neq 0$.

Solution by J. Denmead Smith, The College, Winchester, England. Define $g(x) = [f(xr) - f(r)]/(x-1)$. Thus $\lim_{x \to 1} g(x) = rf'(r)$. Now, for $0 < x < 1$,

$$g(x) = \sum_{n=0}^{\infty} [f(r) - s_n(r)]x^n > [f(r) - s_m(r)]x^m,$$

where m is arbitrary. On letting $x \uparrow 1$ we obtain $rf'(r) \geqq f(r) - s_m(r) > 0$, and hence $f'(r) \neq 0$. It is clearly sufficient to assume that $s_n(r) \geqq f(r)$ for all values of n, and $s_n(r) > f(r)$ for at least one value.

Also solved by Miroslav Ašić (Yugoslavia), Günter Bach (Germany), Robert Breusch (New Zealand), Paul Bruckman, Ron Evans, A. A. Jagers (Netherlands), Yakar Kannai (Israel), Joel Levy, O. P. Lossers (Netherlands), L. E. Mattics, J. G. Mauldon, Michael Skalsky, Paul Smith (Canada), William Stockwell, John Swetits, Joseph Ullman, Robert Vermes (Canada), Man Wah Wong (Canada), and the proposer.

Central Idempotents in a Power Series Ring

6039 [1975, 671]. *Proposed by Robert Gilmer, Florida State University*

Let R be an associative ring and let $\{X_i\}_1^n$ be a finite set of commuting indeterminates over R. Prove that each central idempotent of the power series ring $R[[X_1, \ldots, X_n]]$ is in R.

Solution by J. Brewer and E. Rutter, University of Kansas. A proof of the problem is given in Brewer and Rutter, *Isomorphic polynomial rings*, Archiv. der Math. 23 (1972), 484–488, Lemma 3. To wit, suppose that g is a central idempotent of $R[[X_1, \ldots, X_n]]$. If e is the constant term of g, clearly e is idempotent. If $g \neq e$, let $a \cdot X_1^{k_1} \cdots X_n^{k_n}$ be a term of g having minimal total degree. (Here we assume that g has been written as a formal sum of distinct monomials.) The coefficient of $X_1^{k_1} \cdots X_n^{k_n}$ in g^2 is $ea + ae = 2ae$ and since $g^2 = g$, $2ae = a$. Thus $2ae^2 = 2ae = ae$ and it follows that $ae = 0$. But then $2ae = a = 0$.

Also solved by Mrs. D. Alamelu (India), D. D. Anderson, Douglas Costa, Francis Flanigan, Ralph Grimaldi, Joseph Gruendler, David Handelman (Canada), A. A. Jagers (Netherlands), Peter Landweber, David Lantz, Miguel Laplaza (Puerto Rico), Henry Lieberman, J. G. Mauldon, John O'Neill, Barbara Osofsky, Claude Schochet, Earl Taft, and the proposer.

Integral of a Jacobian

6040 [1975, 672]. *Proposed by Jan Mycielski, University of Colorado*

Let f be a continuously differentiable map of the unit cube I^n into the Euclidean space R^n which

maps the boundary of I^n into one point. Let $J(f, x)$ be the Jacobian determinant of f at x. Prove

$$\int_{I^n} J(f, x)dx = 0.$$

Solution by Thad Dankel, Jr., University of North Carolina. We use the terminology and notation of *Calculus on Manifolds* by M. Spivak (New York, 1965).

Consider the singular n-cube $f:(x^1,\ldots,x^n)\to(y^1,\ldots,y^n)$. Let $\omega = dy^1\wedge\cdots\wedge dy^n$ and $\alpha = y^1dy^2\wedge\cdots\wedge dy^n$, so that $d\alpha = \omega$. The form $f^*\omega$ is given by $f^*\omega = J(f,x)dx^1\wedge\cdots\wedge dx^n$. Thus

$$\int_{I^n} J(f,x) = \int_{I^n} f^*\omega = \int_f \omega = \int_f d\alpha.$$

Now, since f is constant on the boundary of I^n, the faces $f_{(i,j)}$ of f are constant maps, so that $f^*_{(i,j)}\alpha = 0$. Hence

$$\int_{\partial f} \alpha = \sum_{i=1}^{n}\sum_{j=0}^{1}\int_{f_{(i,j)}} \alpha = \sum_{i=1}^{n}\sum_{j=0}^{1}\int_{I^{n-1}} f^*_{(i,j)}\alpha = 0.$$

Using Stokes' Theorem, we get

$$\int_{I^n} J(f,x) = \int_f d\alpha = \int_{\partial f}\alpha = 0.$$

Also solved by George Crofts, Nathaniel Grossman, Yakar Kannai (Israel), Tom Koornwinder (Netherlands), O. P. Lossers (Netherlands), L. E. Mattics, Denmead Smith (England), and the proposer.

A Random Horse Race

6041 [1975, 672]. *Proposed by S. W. Golomb, University of Southern California*

There are n horses in a "random" horse race, in which all $n!$ orders of finish are equally probable a priori. A gambler is allowed to select k horses, to finish first, second,..., kth.

1. What is the probability $Q_n(k, i)$ that exactly i of his k selections will finish among the first k?
2. What is the probability $P_n(k, i)$ that exactly i of his k selections will finish in the precise positions predicted for them?

Solution by Thomas Spencer, Trenton State College. (1) The probability $Q_n(k, i)$ can be computed by direct enumeration as follows. There are $\binom{k}{i}$ ways of choosing the i selections to finish among the first k and $\binom{n-k}{k-i}$ ways of choosing the remainder of the first k, also the first k can be arranged in $k!$ ways and the last $n-k$ can be arranged in $(n-k)!$ ways. Thus the required probability is given by

$$Q_n(k,i) = \frac{\binom{k}{i}\binom{n-k}{k-i}k!\,(n-k)!}{n!} = \frac{\binom{k}{i}\binom{n-k}{k-i}}{\binom{n}{k}}.$$

(2) The probability $P_n(k, i)$ can be computed using a standard result from probability theory (See e.g. William Feller, *An Introduction to Probability Theory and Its Applications*, p. 106) which says: If $A_1,\ldots, A_k$ are k events, the probability that exactly i among the k events occur is given by

$$S_i - \binom{i+1}{i}S_{i+1} + \binom{i+2}{i}S_{i+2} - + \cdots \pm \binom{k}{i}S_k,$$

where $S_0 = 1$ and $S_r = \Sigma P(A_{j_1}A_{j_2}\cdots A_{j_r})$, $1 \leqq r \leqq k$, the sum taken so that each combination occurs

exactly once. Let A_j be the event that the jth horse picked finishes in the jth position, and note that

$$P(A_{j_1}\cdots A_{j_r})=\frac{(n-r)!}{n!}$$

and that there are $\binom{k}{r}$ such terms. The required probability is then given by

$$P_n(k,i)=\sum_{r=i}^{k}(-1)^{r-i}\binom{r}{i}\binom{k}{r}\frac{(n-r)!}{n!}.$$

Also solved by R. L. Andrews, Clarles Barnaby, David Bienenfield (Israel), A. L. Bosch (Netherlands), Benedict Carlat, L. E. Clarke (England), Albert Currier, Robert Douglas, Ralph Grimaldi, Thomas Hern, J. D. Hiscocks, M. Kothmann, R. J. Kulperger (Canada), Harry Lass, Joel Levy, Carl McCarty, Craig Munshower, R. M. Norton, George Pfeiffer, Louis Salkind, Martin Schechter & Peter Borwein (Canada), Michael Skalsky, Wolfe Snow, John Williams (Australia), and the proposer.

Editor's note. The proposer with his solution offered the following observations:

(1) $P_n(n,0)=P_n(n-1,1)=P_n(n,1)-(-1)^n/n!$,

(2) For fixed i, $\lim_{n\to\infty}P_n(n,i)=1/e\cdot i!$ for $i\geqq 0$, which is the Poisson distribution with parameter $\lambda=1$. Is there then a sense in which this is the probability of i fixed points when a "random permutation" is performed on the positive integers?

C^∞ Functions Vanishing Outside [0, 1]

6042 [1975, 766]. *Proposed by F. T. Laseau, G. M. Leibowitz, C. H. Rasmussen and S. J. Sidney, University of Connecticut*

Is every C^∞ real-valued function on the line which vanishes outside $[0,1]$ expressible as a difference of two such functions which are non-negative?

Solution by Jan Boman, University of Stockholm, Sweden. Let f be the given function. It is enough to find a function $h\in C^\infty$ which is non-negative, majorizes f, and whose support is contained in $[0,1]$. For then the solution to our problem is given by $f=h-(h-f)$. The problem to find such a function is trivial except in neighborhoods of the points 0 and 1. It is enough to consider a neighborhood of one of these points, say $x=0$. Set

$$c_k=\sup\{|f(x)|;0\leqq x\leqq 2^{-k}\},\qquad k=1,2,\ldots.$$

Take a non-negative function $u\in C^\infty$, such that $u(x)=0$ for $x<0$ and $u(x)=1$ for $x>1$, and set

$$h(x)=\sum_{k=1}^{\infty}c_k u(2^{k+1}x).$$

Then h is constant for $\frac14\leqq x\leqq\frac12$. Also $f\leqq h$ for $x\leqq\frac12$, for if $2^{-k-1}\leqq x\leqq 2^{-k}$ we have $f\leqq c_k$ and $u(2^{k+1}x)=1$; hence $c_k\leqq h$. To see that $h\in C^\infty$ it is enough to prove (Weierstrass M-test) that

$$(*)\qquad \sum_{k=1}^{\infty}2^{nk}c_k<\infty\quad\text{for every } n.$$

Applying Taylor's formula with Lagrange's remainder to f, we get $f(x)=(x^n/n!)f^{(n)}(\xi)$, where $0<\xi<x$, hence

$$c_k=\sup_{0\leqq x\leqq 2^{-k}}|f(x)|=\sup_{0\leqq x\leqq 2^{-k}}\left|\frac{x^n}{n!}f^{(n)}(\xi)\right|\leqq\frac{2^{-kn}M_n}{n!},$$

where $M_n=\sup|f^{(n)}|$. Hence $2^{kn}c_k$ is bounded for each n, which clearly implies (*). This completes the proof.

Also solved by A. A. Balkema (Holland), Thøger Bang (Denmark), R. A. Christiansen, Edward Howorka, Itrel Monroe, and M. J. Pelling (Nigeria).

Editor's Note. Boman also proves the generalization: Let $f \in C^m$ (m a positive integer) from R^n to R with compact support K. Then there exist non-negative C^m functions g and h with supports contained in K such that $f = h - g$. If $f \in C^\infty$, one may take g and h in C^∞.

Degrees of Irreducible Polynomials over a Field

6043 [1975, 766]. *Proposed by Brian Peterson, University of California, Berkeley*

Let P be a nonempty proper subset of the primes. Consider algebraic extensions F of the rationals Q with the property

(*) Every x in F has degree over Q divisible only by primes in P.

A Zorn's lemma argument shows that there exist maximal extensions satisfying (*). Is such a maximal extension unique up to isomorphism?

Solution by Robert Gilmer, Florida State University. A maximal extension need not be unique up to isomorphism; there is an example to prove this in which $P = \{5\}$. The idea of the example is to produce irreducible quintics $f(X)$, $g(X)$ in $Q[X]$ such that $g(X)$ factors over $Q(u)$, where u is a root of $f(X)$, as the product of an irreducible quadratic and an irreducible cubic. Given such polynomials $f(X)$ and $g(X)$, let F_1 be a maximal extension of $Q(u)$ with property (*) and let F_2 be similarly defined with respect to $Q(v)$, where v is a root of $g(X)$. If F_1 and F_2 were isomorphic, then F_1 would contain a root of $g(X)$, and hence a subfield of degree 10 or 15 over Q, contrary to the assumption that F_1 has property (*).

To produce polynomials $f(X)$ and $g(X)$, let $g(X) = X^5 - n$, where $n = 2^4(5 \cdot 19^2 - 2^{11})$, let v be the real fifth root of $g(X)$, and let $u_1 = (2^3 \cdot 5^3 + 5^3 v)^{1/2}$. A direct computation shows that $u_1 \notin Q(v)$ and hence $[Q(v, u_1): Q] = 10$. Another computation shows that $u = 10u_1$ is a root of the polynomial $-f(X)f(-X) \in Q[X]$, where $f(X) = X^5 - 5aX^3 - rX^2 + tX - m$, $a = 2^4 \cdot 5^5$, $r = -2^6 \cdot 5^8$, $t = 2^7 \cdot 3 \cdot 5^{11}$, and $m = 2^7 \cdot 5^{13} \cdot 19$. Moreover, $f(X)$ and $-f(-X)$ are irreducible in $Q[X]$ since $f(X)$, reduced modulo 3 (which is $X^5 - X^3 - X^2 - 1$), is easily seen to be irreducible. We assume without loss of generality that u is a root of $f(x)$. Then $[Q(u, v): Q(u)] = 2$ implies that $g(X)$ has an irreducible factor of degree 2 over $Q(u)$; moreover, since $Q(u)$ is a real field and v, the only real root of $g(X)$, is not in $Q(u)$, it follows that $g(X)$ decomposes over $Q(u)$ as the product of an irreducible quadratic and an irreducible cubic.

Also solved by John H. Smith, and by the proposer.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN
with the assistance of the mathematics departments of St. Olaf and Carleton Colleges
COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, CARLETON COLLEGE

We invite readers to submit reviews of significant recent college-level mathematics books. We especially encourage reviews based on classroom use, or comparative reviews of several related books. Reviews should ordinarily not exceed two pages (per book) typed double spaced. Manuscripts of reviews as well as books submitted for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057.

Introduction to Applied Numerical Analysis. By Richard W. Hamming. McGraw-Hill, New York, 1971. x + 331 pp. \$16.50.

Numerical Methods with FORTRAN IV Case Studies. By William S. Dorn and Daniel D. McCracken. Wiley, New York, 1972. x + 447 pp. $16.75. (Telegraphic Review, February 1973.)

Numerical Quadrature and Solution of Ordinary Differential Equations. By A. H. Stroud. Springer–Verlag, New York, 1974. xi + 338 pp. $9.50 (paperback). (Telegraphic Review, February 1975.)

Compared to many of the competing texts, each of these books is distinctive, in both the intended audience and the method used to present the basic ideas of numerical analysis. I used each once in the last three years, in the order listed above: Hamming's and Stroud's texts, in a course for upper-division mathematics majors; Dorn and McCracken's, in a course for non-majors at about the sophomore level.

My evaluation of these texts is closely related to my opinions about teaching numerical analysis to liberal arts undergraduates. These opinions, which have evolved from my teaching three survey courses before using Stroud's text, are outlined further in my comments at the end of this review. The basic ideas which I believe should be taught in an introductory numerical analysis course are finite precision computation as a source of error; the estimation and control of roundoff and truncation error; the need for theoretical analysis of the behavior of an algorithm, including stability; and the examination of sufficient conditions for convergence of an algorithm. In addition, I believe that students should learn how to choose among several competing methods for solving a problem. To develop these ideas in an introductory course, I look for a text with geometric motivation where appropriate; concrete analysis of both the theoretical behavior of an algorithm, and its behavior in carefully-chosen examples illustrating important points; sample numerical results illustrating the range of behavior of an algorithm; and good problems at a variety of levels of difficulty.

I have found that most mathematics majors have substantial difficulty in analyzing the error in a method, and in applying the standard "tricks" from analysis used to make a problem more tractable to a numerical solution. Therefore, I believe it wise to choose a definitely elementary introduction to numerical analysis, at least for the first part of the course.

Of these three texts, Dorn and McCracken's comes closest to meeting my criteria for a good introductory text for mathematics students, even though the authors state that it is written for engineers and scientists. Dorn and McCracken present the basic ideas of numerical analysis through an extensive study, at an elementary level, of the standard topics for an introductory course: roots of nonlinear equations, systems of linear equations, numerical integration and differentiation, interpolation, solution of the initial value problem for ordinary differential equations (IVP), and error and its control, plus least squares approximation. They introduce each topic with geometric motivation, prove the basic results analytically, give simple examples with numerical results, and develop substantial "case studies" to illustrate a number of the techniques. There are 13 of these case studies, each illustrating the solution of a "practical" problem. The examples in this text are uniformly excellent: they are as simple as possible, yet clearly illustrate the desired numerical phenomena. The material on roundoff error and numerical instability is especially good. The authors make the analysis of roundoff error straightforward by introducing process graphs and using them to analyze the roundoff error in many of the algorithms. They present numerical instability *via* four case studies, including catastrophic cancellation in the evaluation of Taylor series, and in the evaluation of π by Archimedes' method. There is a flowchart and FORTRAN program with numerical output for each case study and nearly every algorithm. Each chapter concludes with a substantial set of problems, both theoretical proofs and numerical computations, ranging from straightforward exercises (for many of which there are answers) to substantial problems suitable for a term project. Finally, the authors give a valuable annotated list of 25 references, including both introductory and advanced texts, besides the references to the research literature scattered through the text. Unfortunately, there is a substantial

number of misprints. This is a sound introductory treatment of numerical analysis at a more elementary level than most of the competing texts.

Dorn has told me that, despite the elementary nature of his text, it is predominantly used in upper division courses for mathematics majors.

Hamming's text is designed to introduce engineers and scientists to the basic ideas of numerical analysis through a brief survey of many (twelve) topics at an elementary level, including optimization, Fourier series, and random processes. Each is treated "as simply as possible and in a uniform way". The algorithms are explained geometrically and often illustrated by a simple numerical example, sometimes showing the limitations of the algorithm. Since the emphasis is on basic ideas, rather than computer implementation, there are no computer programs for the algorithms, and only a few flowcharts (for solving one nonlinear equation, the first problem discussed). The exercises are consistent with Hamming's philosophy and level of presentation: there are very few compared to most texts, and most are designed to show the student how a method performs on different problems by a direct application of the material just covered. There are references to advanced or research level literature scattered through the text, as well as substantial doses of the author's well-known and distinctive philosophy toward numerical analysis. Although Hamming is successful in introducing the reader to the main ideas of a wide variety of methods, I think most mathematics instructors would agree that this text is too cursory for use in a mathematics course.

Stroud writes for mathematicians at the senior or beginning graduate level. He gives a theoretical and rigorous treatment of quadrature (in one variable) and the IVP, plus the necessary background in interpolation. The book is written in a rather formal, primarily theorem-proof style, with minimal geometrical motivation for an introductory text. There are FORTRAN programs, some with tables of numerical results, for the more complicated algorithms, but few examples compared to the other two books. Surprisingly, Stroud doesn't emphasize the key role of linearity: he does not state that the Lagrange interpolation polynomials form a basis; he uses ideas concerning linear functionals to derive the Peano error estimates for quadrature rules, but does not emphasize the role of linearity. There is very little material on roundoff error and its analysis; nothing on adaptive quadrature, or the choice among competing methods (aside from references); and few applications. There are a reasonable number of exercises, but very few routine problems to build students' confidence. The text includes optional sections on such topics as spline interpolation and rational extrapolation, and each chapter concludes with a substantial list of references to advanced and research literature. There are a number of misprints, and I would be happy to supply a list of those I found. I noted only one serious error, in the last subscript on r in the third equation in (1), page 232. These equations represent the motion of a satellite in a restricted three-body problem. We found that the equations as printed do not have solutions which are closed orbits. Stroud has written a solid text, especially on quadrature, where he gives a number of results on convergence and error estimation not commonly found in an introductory text. However, I feel this text is too abstract and difficult for an introductory course for the average mathematics major.

These different approaches are nicely illustrated by each text's treatment of predictor-corrector (P − C) methods for solving the IVP. Hamming uses direction fields to motivate Euler's method, then introduces P − C methods by a method of order two. Using truncated Taylor's series, he obtains error estimates, which he applies to decide when to change step size, and gives one example using this method. There are five exercises, all involving hand computation, and several graphs to give a geometric interpretation of the development. Stroud introduces general predictor methods using Newton-Cotes quadrature, including numerical results using a predictor alone compared to using a P − C pair. After introducing stability by studying predictor methods applied to $y' = \lambda y$ and proving several theorems, he treats a relatively simple example in considerable detail, formally defines stability, and gives tables indicating the regions of stability for a number of predictors. There are seven exercises for this material, six of which ask the student to solve an IVP, presumably by writing a

computer program. Dorn and McCracken use geometric representations to give a very clear exposition of a $P-C$ method of order two, and derive a sufficient condition for being able to iterate the corrector to convergence. They obtain the truncation error by using Taylor's series (in one variable) with remainder, use this error estimate to discuss the choice of step size, and give an elementary discussion of stability. Finally, they discuss a substantial example (a "case study"), large deflections of a beam, and compare Runge-Kutta and $P-C$ methods. There are 22 exercises on $P-C$ methods, quite a few of which extend the ideas in the text.

My small sample of students (there were two or three in each class) agreed with the evaluations I have given. My students found Stroud's presentation of differential equations, especially stability, rough going, as it seems to be substantially more difficult than the preceding chapters. They also became discouraged in working problems, as they really needed more routine problems for building confidence. The two rather average sophomores who used Dorn and McCracken's text found the book readable, and the exercises appropriate. Each course was intensive: about two and one-half hours of class per day, five days per week, for 18 days.

I have a strong, though unorthodox, opinion of the role of numerical analysis in the mathematics curriculum in a liberal arts college. First, I believe it is much more important to teach a thorough understanding of the basic ideas of numerical analysis — the development and analysis of good algorithms, and their use in solving real problems — than to give a survey of a large number of standard techniques. Second, I believe that the material should be related to and develop further ideas the student has seen in previous courses, and develop the student's ability to carry out error estimation and compare numerical methods. After teaching three survey courses, I concluded that, especially in our intensive academic schedule, there is not enough time to attain these objectives in such a course. The topics in Stroud's text are one obvious choice for what I believe is a better alternative: studying several related topics in depth. A second possibility would be the study of numerical linear algebra, but I know of no text on this subject suitable for the average undergraduate.

I found that an intensive study of quadrature and the solution of the IVP is a good alternative to the standard introductory survey course. This material is easily accessible to the average undergraduate, and in presenting it the instructor has ample opportunities to develop the basic principles of numerical analysis. I know of no better way to achieve the second goal stated in the previous paragraph than the evaluation of integrals by various methods, especially improper integrals. In addition, the estimation of the difference between a given improper integral and one which can be evaluated by hand or by a numerical method would be a good way to introduce a student to the ideas of error estimation, an idea with which my students have had difficulty.

Dorn and McCracken provide a fine readable introduction to quadrature and the solution of the IVP, although the material on quadrature is rather brief. I believe that a student must see the kind of examples they give in order to really understand the basic ideas of numerical analysis. In addition, their analytical development, in the text and in problems, demonstrates the type of theoretical development one must do to really understand a numerical method. For the theoretical development of these methods which an upper division student should see, their text should be supplemented with a book such as Wendroff's *First Principles of Numerical Analysis* (Addison-Wesley, 1969; reviewed in this *Monthly*, September 1970, pp. 788–789). Good problems on quadrature can be found in the text by Acton, and that by Shampine and Allen (TR January 1971, p. 104, and December 1973, p. 1160, respectively). I intend to use such a combination next time I teach numerical analysis; I believe it would form an excellent introduction.

JOHN KARON, Colorado College

TELEGRAPHIC REVIEWS

Telegraphic reviews are designed to give prompt notice of new books with sufficient information to assist our readers in deciding whether to order an examination copy or to suggest library purchase. Possible uses are indicated as follows:

T = textbook P = professional reading
S = supplementary reading L = undergraduate library purchase
13 to 18 = freshman to second year graduate level usage
1 to 4 = appropriate time in semesters to cover text

Asterisks (*) or question marks (?) denote special positive or negative emphasis, respectively. Publishers are denoted by standard abbreviations; complete addresses may be found in *Books in Print*.

GENERAL, S*, P*, L. *Les Objets Fractals: Forme, Hasard et Dimension.* Benoit Mandelbrot. Flammarion, 1975, 190 pp, (P). An absolutely unique interdisciplinary exploration of "self-similar" patterns in nature and in mathematics: snowflake curves, shorelines, distribution of stars, holes in Swiss cheese, word frequencies, family incomes. Each has the property that when suitably magnified, the resulting pattern is unchanged, except in density or measure. Mandelbrot discusses various ways of calculating these measures, drawing on classical works in real analysis, probability, and Brownian motion, as well as on more speculative insights of the philologist George Zipf after whom these patterns are sometimes named. The measures discussed by Mandelbrot extend the notion of dimension to non-integral values, hence the name of the book. LAS

GENERAL, S(9-16). *Logik Macht Spass: 85 Aufgaben mit Lösungen.* György Bizám, János Herczeg. Akadémiai Kiadó, 1976, 391 pp, $16. Extended and clever treatment of deductive problems of the Smith-Jones-Robinson, brakeman-fireman-engineer-type and its generalizations, relying for solutions on incidence tables in two and three dimensions. The problems (and extensively annotated solutions) are carefully sequenced to build on each other, though the authors have marked several short trails for those who cannot make the entire journey at once. No mathematical prerequisites whatever. PJC

GENERAL, L. *Index of Mathematical Papers, V. 6.* AMS, 1976. *Part 1:* 666 pp; *Part 2:* 500 pp, $80 (P). Index to the 1974 Volumes 47 and 48 of *Mathematical Reviews*. Part 1 is a subject index; Part 2 is the author index. LAS

PRECALCULUS, T(13-14), S. *College Algebra and Trigonometry.* William G. Ambrose. Macmillan, 1977, xi + 547 pp, $11.95. Provides precise definitions for all concepts used and proofs for theorems presented. Many examples illustrate the abstractions. Very large number of section and chapter review problems immediately followed by answers. Convenient summary of all formulas. Tables. Index.RJA

PRECALCULUS, T(13-14: 1), S. *Elementary Functions.* Carl B. Allendoerfer, Cletus O. Oakley, Donald R. Kerr, Jr. McGraw, 1977, xii + 314 pp, $12.50. Text built around concept of a function and specific examples of functions: polynomial, exponential, logarithmic, and trigonometric. Investigates properties of each. Detailed explanations for each example. Could be used in a terminal course. Answers to odd-numbered problems. Tables. Index. RJA

PRECALCULUS, T(13: 1). *Pre-Calculus Mathematics, Second Edition.* Hal G. Moore. Wiley, 1977, x + 517 pp, $13.95. The main change from the first edition (TR, October 1973) is an expansion of the first chapter, Fundamentals, into two chapters, Fundamentals and Functions, with added material in each. LLK

PRECALCULUS, T(13: 1). *Plane Trigonometry, Seventh Edition.* Fred W. Sparks, Paul K. Rees, Charles Sparks Rees. P-H, 1977, xi + 316 pp, $10.95. Another edition of a very successful trigonometry book (Sixth edition TR, February 1972). Many sections have been rewritten and topics rearranged, but it remains a classical, useable text. LLK

NUMBER THEORY, P. *Lecture Notes in Mathematics-546: Lectures on Transcendental Numbers.* Kurt Mahler. Springer-Verlag, 1976, xxi + 254 pp, $10.20 (P). This book is derived from lectures given by Mahler during the last twenty years and is edited and sometimes enlarged by Diviš and LeVeque. It aims at giving accounts of some classical results and in particular to present in detail the Siegel-Shidlovski theory. CEC

LINEAR ALGEBRA, T(14: 2). *An Introduction to Linear Algebra for Science and Engineering, Second Edition.* Dominic G.B. Edelen, Anastasios D. Kydoniefs. Am Elsev, 1976, xvii + 269 pp, $10.95. Revisions from the first edition (TR, January 1973), in addition to corrected errors and misprints, include a lecture on least squares and additional numerical examples and problems. LLK

LINEAR ALGEBRA, T(14: 1). *Elementary Linear Algebra, Second Edition.* Bernard Kolman. Macmillan, 1977, xii + 314 pp, $10.95. This edition differs from the first edition (TR, June 1970, January 1971; ER, March 1974) in the unification of inner product spaces with real vector spaces and of the material on eigenvalues and eigenvectors. There are additional examples and exercises and more geometric material. LLK

ALGEBRA, P. *Lecture Notes in Mathematics-530: Weil's Representation and the Spectrum of the Metaplectic Group.* Stephen S. Gelbart. Springer-Verlag, 1976, 140 pp, $7.40 (P). A study of automorphic forms on the metaplectic group. PJM

ALGEBRA, S(16-17). *Algèbre, Solutions Développées des exercises.* S. MacLane, G. Birkhoff. Gauthier-Villars (US Distr: SMPF, 111 W. 57th St., New York 10019), 1976, v + 184 pp, 48 F (P). Solutions for chapters 10, 13, 16 and 17 of the French edition of MacLane and Birkhoff's *Algebra*. Covers finite abelian groups, structure of groups, multilinear algebra and Galois theory. The French is relatively easy, particularly if a copy of MacLane-Birkhoff is available. PJM

FINITE MATHEMATICS, T(13-14: 1). *Finite Mathematics With Applications in Business, Biology, and Behavioral Sciences.* Margaret L. Lial, Charles D. Miller. Scott F, 1977, 434 pp, $12.95. An attractive feature is the collection of 17 cases applying content of the sections to real-world situations, e.g., bidding on a potential oil field, linear programming in water pollution. Includes selection of questions from CPA exams. No computer programming. Skimpy on exercises. Some portions taken from *Mathematics: With Applications in the Management, Natural, and Social Sciences,* by the same authors. JG

CALCULUS, T(13: 1). *Mathematics for Financial Analysis.* Michael Gartenberg, Barry Shaw. Pergamon Pr, 1976, x + 210 pp, $12. This is a short course in calculus with applications to business and economics. It also includes a chapter on vectors and matrices and one on linear programming. LLK

COMPLEX ANALYSIS, T(16-18: 2, 3), L. *Several Complex Variables.* H. Grauert, K. Fritzsche. Grad. Texts in Math., V. 38. Springer-Verlag, 1976, viii + 207 pp, $18.80. A modern, algebraic approach to the subject. Chapters on holomorphic functions, domains of holomorphy, the Weierstrass preparation theorem, sheaf theory, complex manifolds, cohomology and real methods. No exercises, but many examples. References to books (as opposed to individual papers). PJM

DIFFERENTIAL EQUATIONS, S(18), P. *Partial Differential Equations in the Complex Domain.* D.L. Colton. Pitman, 1976, 88 pp, $3.90 (P). An introduction to the use of function theoretic methods in the investigation of a class of physically motivated problems which are improperly posed initial value problems. These lecture notes are devoted to developing an analytic theory of partial differential equations based on the analytic theory of functions of a complex variable. CEC

DIFFERENTIAL EQUATIONS, S(14-15). *Introduction to Qualitative Theory of Differential Equations.* J. Plante. U of No Carolina, 1976, 70 pp, $4 (P). A brief introduction to qualitative theory covering Poincaré-Bendixson, conservative and dissipative systems, stability, and some non-linear second order equations. Accessible to good undergraduates. No exercises. SG

FUNCTIONAL ANALYSIS, T(18: 1, 2), S, P. *Vorlesungen über nichtlineare Operatorengleichungen.* Thomas Riedrich. B.G. Teubner, 1976, 182 pp, 16 M (P). A thorough exposition of the basic facts required for understanding the present state of nonlinear functional analysis. Presents many examples concerning existence theorems (finite and infinite dimensional), Leray-Schauder theory, mapping degree, simple approximation methods and perturbations of nonlinear operational equations. JAS

OPTIMIZATION, P. *Lecture Notes in Economics and Mathematical Systems-117: Optimization and Operations Research.* Ed: W. Oettli, K. Ritter. Springer-Verlag, 1976, iv + 316 pp, $11.50 (P). Proceedings of the conference held at Oberwolfach July 27 to August 2, 1975. JAS

ANALYSIS, P. *Fixed Point Theory and Its Applications.* Ed: Srinivasa Swaminathan. Acad Pr, 1976, xiii + 216 pp, $12. Contains the proceedings of a seminar held at Dalhousie University, June 9-12, 1975. Twenty-one papers with references. List of participants. Index. RJA

DIFFERENTIAL GEOMETRY, T(16-17), L. *Global Theory of Connections and Holonomy Groups.* André Lichnerowicz. Trans. and Ed: Michael Cole. Noordhoff, 1976, xiv + 250 pp, Dfl. 65. Translated from the French original. The bibliography is a little old (up to mid-fifties), but addition of subject and symbol indices and topological changes are worthwhile additions to the translation. Starting with the notion of differentiable manifolds, but assuming some knowledge of geometry, the author develops the global theory of connections. PJM

GEOMETRY, T(17-18), P. *Integral Geometry and Geometric Probability.* Luis A. Santaló. A-W, 1976, xvii + 404 pp, $19.50. A thorough and engaging survey, the first volume in the planned Encyclopedia of Mathematics and Its Applications edited by G.C. Rota. Discusses integral geometry in the plane, in higher dimensional Euclidean spaces, and in spaces of constant curvature. Contains extensive notes and a number of exercises. Experts and nonexperts will benefit from this excellent book. SG

GEOMETRY, P. *Vorlesungen über Geometrie der Algebren.* Walter Benz. Grund. math. Wissenschaften, B. 197. Springer-Verlag, 1973, xi + 368 pp, $32.60. Axiomatic treatment of many recent developments in geometric algebra. Generalizations of Möbius-, Laguerre-, Lie-, and Minkowskian-geometries. Emphasis is far more geometric than most studies of geometric algebra. SS

TOPOLOGY, P. *The Topology of Stiefel Manifolds.* I.M. James. London Math. Soc. Lect. Notes, No. 24. Cambridge U Pr, 1976, viii + 168 pp, $7.95 (P). An updating of 1961 course lectures given at Harvard. Prerequisites include algebraic topology at the level of Spanier's text. A list of research problems is included. JAS

TOPOLOGY, T(15-17: 1), S, P, L. *Differential Topology with a View to Applications.* D.R.J. Chillingworth. Pitman, 1976, 291 pp, $17 (P). A gentle introduction designed to make the methods of differential topology accessible to applied scientists. Assumes only calculus and elementary linear algebra; emphasizes the global qualitative behavior of dynamical systems. LAS

TOPOLOGY, S(17), P, L. *Lectures on Hilbert Cube Manifolds.* T.A. Chapman. CBMS Reg. Conf. in Math., No. 28. AMS, 1976, x + 131 pp, $7.20 (P). The countable product of copies of [-1,1] is called the Hilbert cube. A Hilbert-cube manifold is a separable metric manifold modelled on the Hilbert cube. These lectures from a 1975 regional conference survey some of the results in the theory of Hilbert-cube manifolds and also includes a list of open problems. PJM

TOPOLOGY, P. *On PL de Rham Theory and Rational Homotopy Type.* A.K. Bousfield, V.K.A.M. Gugenheim. Memoirs No. 179. AMS, 1976, ix + 94 pp, $7.20 (P). A description of de Rham cohomology for PL-manifolds using Quillen's homotopical algebra. PJM

TOPOLOGY, P. *Lecture Notes in Mathematics-533; The Homology of Iterated Loop Spaces.* Frederick R. Cohen, Thomas J. Lada, J. Peter May. Springer-Verlag, 1976, vii + 490 pp, $15.20 (P). A collection of four papers which summarize and complete the study of homology operations and of their application to "the calculation of, and analysis of internal structure in, the homologies of various spaces" of current interest. A fifth develops an up-to-homotopy notion of an algebra over a monad with applications to iterated loop spaces. JAS

PROBABILITY, P. *Stochastic Processes with Learning Properties.* Sándor Csibi. Springer-Verlag, 1975, 150 pp, $11.10 (P). Contains materials given in two short courses at Udine, Italy entitled "Simple and Compound Processes in Machine Learning" and "Stability and Complexity of Learning Processes." Both deal with convergence of algorithms which possess a "learning" capability. Of particular interest to control theorists with a probabilistic inclination. TAV

PROBABILITY, T(18: 1), P, L. *Lecture Notes in Mathematics-529: Doubly Stochastic Poisson Processes.* Jan Grandell. Springer-Verlag, 1976, x + 234 pp, $10.30 (P). In the typical Poisson process λ is a constant or known function of time. The author considers the properties of processes where λ is a random variable, possibly time dependent. Knowledge of measure theory and Hilbert space is helpful, but not essential. A variety of applications are developed. An interesting treatment. TAV

STATISTICS, T?(13). *Statistics for the Social Scientist: 1. Introducing Statistics.* K.A. Yeomans. Penguin, 1974, 259 pp, $7.95 (P). A reprint of the 1968 edition. Volume 1 deals with graphical and numerical methods at a precalculus level. The treatment is essentially descriptive with applications appearing in Volume 2. The catch: Volume 2 is no longer available! TAV

STATISTICS, T(16-17: 1), L. *Introduction to Nonparametric Detection with Applications.* Jerry D. Gibson, James L. Melsa. Math. in Sci. and Eng., V. 119. Acad Pr, 1975, xii + 241 pp, $24.50. Hypothesis testing aimed at engineering problems that arise in radar, sonar, acoustics, geophysics. Neyman-Pearson approach emphasized. Requires some background in probability. Includes applications, some problems. LH

STATISTICS, T(15-16: 1), S, L. *Modern Factor Analysis, Third Edition Revised.* Harry H. Harman. U of Chicago Pr, 1976, xx + 487 pp, $20. Intended as both a textbook and a state-of-the-art survey. Develops the linear algebra used. Extensive revision of second edition. New material includes canonical, image and alpha factor analysis. Excellent bibliography. LH

STATISTICS, T(16-17). *Tratarea Matematică a Datelor Experimentale.* Ioan Todoran. Editura Academiei România, 1976, 325 pp, Lei 15,50. A substantial exposition of the practical details of correlation theory and appropriate (classical) numerical methods. In Romanian. No exercises. JAS

STATISTICS, T*(13: 1), S*, L. *Statistics, A Beginning.* Roy R. Kuebler, Harry Smith, Jr. Wiley, 1976, xiii + 320 pp, $12.95. An appealing book with excellent examples and problems. Little theory, but concepts are correctly stated and convincingly explained. Basic but not remedial. Graphical and numerical summarization of data, basic probability, confidence intervals, hypothesis testing, contingency tables, regression analysis. Manual for teachers, workbook for students available. RBK

COMPUTER SCIENCE, P. *Algorithms and Complexity, New Directions and Recent Results.* Ed: J.F. Traub. Acad Pr, 1976, ix + 523 pp, $19.50. Proceedings of a symposium held at Carnegie-Mellon University, April 7-9, 1976. Volume contains the texts of 14 invited papers, titles and abstracts of 85 contributed papers, and a list of participants. RJA

COMPUTER SCIENCE, S(13-18), P. *Microcomputer Dictionary and Guide.* Charles J. Sippl, David A. Kidd. Matrix Pub, 1975, ix + 680 pp, $25 (P). Definitions of more than 5,000 terms currently used in micro-electronics fields and seven appendices: Symbols, Units and Constants of Electronics; Mathematics Definitions; Statistics Definitions; Electronics and Computer Acronyms and Abbreviations; Computer Language Summaries; APL, BASIC and FORTRAN; Computer Number and Binary Switching Systems; and Definitions of Programmable Calculator Terms. Handy reference guide, well organized. RJA

COMPUTER SCIENCE, S(16-18), P, L. *Encyclopedia of Computer Science and Technology, V. 1-3.* Eds: Jack Belzer, Albert G. Holzman, Allen Kent. Dekker, 1975, $75 each. *V. 1: Abstract to Amplifiers*, ix + 497 pp; *V. 2: AN/FSQ to Bal*, vii + 494 pp; *V. 3: Ball to Box*, vi + 503 pp. Each topic is treated in a single, comprehensive, encyclopedic article which is self-contained and individually written by a contributing author. References accompany each article. An extremely valuable, well organized reference tool whose only drawback is its high cost. RJA

COMPUTER SCIENCE, T(13-14: 1), S. *Computers, Their Impact and Use: Basic Language.* Robert E. Lynch, John R. Rice. HR&W, 1975, xi + 398 pp, $10 (P). Contains the historical development of computing and its current impact on society. Includes both elementary and advanced topics in the programming language BASIC. Program organization. The material on computer games is especially enticing. A well-done potpourri for the beginner. Problems. References. Appendices. Glossary of technical terms. RJA

COMPUTER SCIENCE, T(15-18: 1), S, P, L. *Operating System Principles.* Per Brinch Hansen. P-H, 1973, xviii + 366 pp, $17.95. Begins with historical development. Proceeds to sequential and concurrent processes, processor and storage management, scheduling and protection. Ends with an in-depth case study of the RC 4000 multiprogramming system. Throughout, algorithms are given in PASCAL. Annotated bibliography for each chapter. Exercises and answers. List of definitions of technical vocabulary. Index to vocabulary. Index to algorithms. Index. RJA

COMPUTER SCIENCE, P. *Pattern Recognition and Artificial Intelligence.* Ed: C.H. Chen. Acad Pr, 1976, ix + 621 pp, $28.50. Twenty-eight papers from a workshop held in June 1976 at Hyannis, Massachusetts. A few other papers were published by IEEE. JAS

COMPUTER SCIENCE, T(14-17: 1), S, L. *Introduction to Programming Languages.* W. Wesley Peterson. P-H, 1974, ix + 358 pp, $14.50. Contains individual presentations on eight programming languages: BASIC, FORTRAN, ALGOL, PL/I, APL, COBOL, SNOBOL, LISP. Discussion of essential features of each language. Includes material on the development and special uses of each language. Comparison of languages through examples programmed in several languages. Appendices. Index. RJA

COMPUTER SCIENCE, T(16-18: 1), S, P, L. *Applied Computation Theory: Analysis, Design, Modeling.* Ed: Raymond T. Yeh. P-H, 1976, xiii + 624 pp, $22.50. Collection of fourteen "tutorial" papers on the fundamentals and applications of computation theory. Editorial introductions to the papers provide background information and emphasize important concepts that are to follow. Appendices. Index. RJA

COMPUTER SCIENCE, S(14), L. *An Introduction to the Uses of Computers.* Murray Laver. Cambridge U Pr, 1976, viii + 232 pp, $6.95 (P). An appreciation of the uses of computers. Intended for the occasional, non-mathematical user. Discusses hardware, software, I/O, programming and applications especially in business and government. Discussion questions with brief answers. Short bibliographies. RWN

SYSTEMS THEORY, P. *Lecture Notes in Economics and Mathematical Systems-115: Foundations of System Theory: Finitary and Infinitary Conditions.* B.D.O. Anderson, M.A. Arbib, E.G. Manes. Springer-Verlag, 1976, viii + 93 pp, $7.40 (P). Applications of category theory to general systems; primarily to finite state machines. PJM

SYSTEMS THEORY, P. *Studies in Operating Systems.* R.M. McKeag, R. Wilson. APIC Stud. in Data. Proc., No. 13. Acad Pr, 1976, x + 263 pp, $21. Case studies of four diverse operating systems: (1) Burroughs B5500 Master Control Program; (2) Titan Supervisor; (3) CDC 6000 Series Scope; (4) The Multiprogramming System. Studies are structured so that easy comparison of features can be made. RJA

SYSTEMS THEORY, S(18), P. *Computer Networks: Text and References for A Tutorial, Revised and Updated.* Marshall Abrams, Robert P. Blanc, Ira W. Cotton. IEEE Computer Society, 1976, iv + 317 pp, $12 (P). Intended for the potential user of large computer communication networks. Takes a "survey" approach rather than providing an in-depth analysis of specifics. Each chapter contains expository material on the topic under consideration followed by selected representative articles by various authors. RJA

APPLICATIONS (BIOLOGY), S(18), P. *Biophysical and Physiological Systems Analysis.* Erol Basar. A-W, 1976, xvii + 429 pp, $12.50 (P); $24.50. This book develops a new methodology combining tools from general systems theory and from bioscience in order to analyze and understand various phenomena in physiology. Utilizes mathematical and statistical techniques such as Laplace transform and time series analysis. Selected examples include the vascular systems, smooth muscles and electrical activity of the brain. I-CH

APPLICATIONS (BIOLOGY), P. *Combinatorial Methods in Developmental Biology.* J.K. Percus. Courant Inst, 1977, iii + 202 pp, $6.25 (P). A well-written survey of biological problems which lead to combinatorial models. Both stochastic and non-stochastic models are considered. Fairly accessible to both biologists and mathematicians. A few exercises are included. SG

APPLICATIONS (BIOLOGY), T(18: 1), S, P, L?. *Partial Differential Equations in Biology.* Charles S. Peskin. New York U, 1976, iv + 227 pp, $7 (P). The electro-physiology of nerves, the inner ear, the retina, and pulse wave propagation in arteries are modeled as partial differential equations. The solutions of these equations are discussed in detail. CEC

APPLICATIONS (CHEMISTRY), P. *Chemical Applications of Graph Theory.* Ed: A.T. Balaban. Acad Pr, 1976, xii + 389 pp, $14.50. Eleven invited chapters survey (for the first time) the broad interface between chemistry and graph theory including applications to molecular structure and identification of isomers (in which graph vertices represent atoms) and to complex reaction processes (in which graph vertices represent reagents). LAS

APPLICATIONS (ENGINEERING), T(16-17: 2), P. *Stability of Fluid Motions.* Daniel D. Joseph. Tracts in Nat. Philo., V. 27-28. Springer-Verlag, 1976, $39.80 each. *V. I*, xiii + 282 pp; *V. II*, xiv + 274 pp. Volume 1 deals with the general theory of stability, instability and bifurcation with applications to various flows. Volume II treats extensions of the theory and applications, emphasizing the role of stability theory in fluid mechanics. The price noted is per volume! Volume II contains an extensive bibliography. TAV

APPLICATIONS (ENGINEERING), P. *Data Compression.* Ed: Lee D. Davisson, Robert M. Gray. DH&R, 1976, xv + 407 pp, $25. "The science of processing information to obtain a simple representation with a tolerable loss of fidelity." Contains forty-six papers on the theory and practice of data compression. The papers are divided into topical areas; each group is preceded by editorial commentary which attempts to provide a common context for that section. Author citation index. Subject index. RJA

Reviewers Whose Initials Appear Above

Richard J. Allen, St. Olaf; Paul J. Campbell, St. Olaf; Clifton E. Corzatt, St. Olaf; Jennifer Galovich, St. Olaf; Steven Galovich, Carleton; Loren Haskins, Carleton; Ih-Ching Hsu, St. Olaf; Roger B. Kirchner, Carleton; Pierre J. Malraison, Carleton; R.W. Nau, Carleton; Seymour Schuster, Carleton; J. Arthur Seebach, Jr., St. Olaf; Lynn A. Steen, St. Olaf; T.A. Vessey, St. Olaf.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, NW, Washington, D.C. 20036. Items must be submitted at least five months before publication can take place.

PERSONAL ITEMS

Professor J. J. Malone, Worcester Polytechnic Institute, represented the Association at the inauguration of Jean Mayer as President of Tufts University on September 18, 1976.

Professor Emeritus Nura D. Turner represented the Association at the inauguration of Father Hugh Francis Hines, O.F.M., as Seventh President of Siena College on October 7, 1976.

Bucknell University: Assistant Professors P. I. Nelson and A. R. Schweinsberg have been promoted to Associate Professors.

Miami University: Assistant Professors Fredric Pollack and Frederick Schuurmann have been promoted to Associate Professors; Associate Professor Shih-Hsiung Tung has been promoted to Professor.

Northwest Missouri State University: Assistant Professor Jo Ingle has been promoted to Associate Professor; Associate Professor David Bahnemann has been promoted to Professor.

Queens College-CUNY: Assistant Professor Robert Cowen has been promoted to Associate Professor; Associate Professor Joseph Hershenov has been promoted to Professor.

State University of New York at Buffalo: Dr. Curtis Greene, MIT, has been appointed Associate Professor; Associate Professor Edith Schneckenburger retired in July 1976 with the title of Emeritus Associate Professor.

Syracuse University: Dr. Barry Burd, University of Illinois, and Dr. Robert Connelly, Jr., Cornell University, have been appointed Assistant Professors; Associate Professor Jack Graver has been promoted to Professor.

University of Connecticut: Dr. L. F. Richardson, Louisiana State University, has been appointed Visiting Associate Professor; Assistant Professor P. R. Fallone, Jr., has been promoted to Associate Professor; Associate Professor Manuel Lerman has been promoted to Professor.

University of Miami: Associate Professor Allan Zame has been promoted to Professor; Dr. Edwin Duda has been appointed Acting Chairman of the Department of Mathematics.

University of Oklahoma: Dr. R. I. Matthews, University of Idaho, has been appointed Instructor; Dr. G. M. Reekie has been appointed Instructor; Ms. Karen H. Chase, Texas A & M University, has been appointed Instructor; Dr. R. M. Schori, Louisiana State University, has been appointed Visiting Professor; Dr. A. R. Sourour, University of Toronto, has been appointed Visiting Assistant Professor.

Associate Professor J. E. Allen, North Texas State University, has been appointed Chairman of the Department of Mathematics.

Assistant Professor of Mathematics Dale Craft, Westmoreland County Community College, has been appointed Assistant Dean of Arts and Science.

Associate Professor H. E. Heatherly, University of Southwestern Louisiana, has been promoted to Professor. He has also been named a Distinguished Professor for 1976–77.

Associate Professor W. M. Page III, Alliance College, has been promoted to Professor.

Assistant Professor W. C. Ramaley, Fort Lewis College, has been promoted to Associate Professor.

Associate Professor Louis Ross, the University of Akron, has been promoted to Professor; Professor Ross has been elected a Fellow of the American Society for Quality Control.

Dr. R. P. Rozek, The University of Iowa, has been appointed Assistant Professor of Economics at the University of Pittsburgh.

Professor Charles L. Seebeck, University of Alabama, retired on July 12, 1976, with the title of Professor Emeritus.

Assistant Professor Frank Servas, Jr., Pratt Institute, has been appointed Assistant Professor of Mathematics and Computer Science at St. John's University, Jamaica, New York.

Professor S. M. Shah, University of Kentucky, retired in May 1976 with the title of Emeritus Professor, after a total teaching service of 45 years (including 10 years at the University of Kentucky). In honor of his retirement the University of Kentucky hosted, with partial support from NSF, a Conference on Complex Analysis from May 19 to May 22, 1976.

Assistant Professor T. H. Slook, Temple University, has been promoted to Professor.

Professor Loren T. Black, Long Beach, California, died in March 1976 at the age of 77. He was a member of the Association for forty-eight years.

Dr. Carl B. Boyer, Brooklyn College-CUNY, died on April 26, 1976, at the age of 69. He was a member of the Association for thirty-nine years.

Professor Emeritus John Wildeboor Hurst, Montana State University, died on September 6, 1976, at the age of 80. He was a member of the Association for forty-eight years.

Professor Torsten Norvig, Wellesley College, died on August 27, 1976, at the age of 50. He was a member of the Association for nineteen years.

Mr. Harold J. Ruth, Ohio University, died on January 25, 1975, at the age of 32. He was a member of the Association for two years.

Dr. Norman P. Salz, Rochester Institute of Technology, died on May 13, 1976, at the age of 55. He was a member of the Association for sixteen years.

Dr. Levi S. Shively, Ball State University, died on July 17, 1976, at the age of 92. He was a Charter Member of the Association.

Professor Emeritus Albert G. Wootton, University of Maine, died on August 17, 1976, at the age of 66. He was a member of the Association for twenty-four years.

GAUSS BICENTENNIAL SYMPOSIUM

At the request of the Royal Society of Canada, a symposium will be held at the Ontario Science Centre, June 3–4, 1977, to celebrate the bicentennial of Gauss' birth. The program will be as follows:

Friday, June 3
- 3:30 P.M. DIEUDONNÉ (Nice)—Algebra & Analysis
- 5:00 P.M. FORBES (Edinburgh)—Astronomy
- 6:30 P.M. Dinner
- 8:00 P.M. MAY (Toronto)—Historical Introduction

Saturday, June 4
- 9:30 A.M. COXETER (Toronto)—Geometry
- 11:00 A.M. GARLAND (Toronto)—Geomagnetism
- 12:30 P.M. Lunch
- 2:30 P.M. SELBERG (Princeton)—Number Theory
- 3:30 P.M. SPROTT (Waterloo)—Statistics
- 5:00 P.M. Reception

Accommodation will be available June 2–5 at Whitney Hall, a residence attached to University College on St. George St. Reservations should be made with the undersigned as soon as possible and no later than May 15, 1977; single rooms $10.00 per night, some double rooms available. Buses will provide transportation to the Science Centre.

G. DE B. ROBINSON, *Chairman of Organizing Committee*

INTERNATIONAL SYMPOSIUM ON CONTINUUM MECHANICS AND PARTIAL DIFFERENTIAL EQUATIONS — INSTITUTO DE MATEMÁTICA, UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

An International Symposium on Continuum Mechanics and Partial Differential Equations will be held at the Instituto de Matemática, Universidade Federal do Rio de Janeiro, on August 1–5, 1977. Further details and registration forms may be obtained from the Symposium Coordinator, Professor L. A. Medeiros, Instituto de Matemática — UFRJ, C. Postal 1835, ZC 00, 20.000 — Rio de Janeiro, Brazil.

CRYPTOLOGIA

CRYPTOLOGIA, a journal devoted to all aspects of cryptology, will be published four times a year. There is a need for a forum for the exchange of ideas related to cryptology in the public sector. This journal will meet that need with research papers, survey articles, personal accounts, reviews, educational notes, and problems. Some of the areas which will be discussed are mathematical, computational, literary, historical, political, military, mechanical, and archeological aspects of cryptology. The Editors of CRYPTOLOGIA are Cipher A. Deavours, Department of Mathematics, Kean College; David Kahn, Department of Journalism, New York University; and Brian J. Winkel, Department of Mathematics, Albion College. Further information may be obtained from: CRYPTOLOGIA, Albion College, Albion, Michigan 49224.

WORKSHOPS SPONSORED BY THE MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA SECTION

The MD-DC-VA section of the MAA will sponsor two five-day workshops June 6–10 and 13–17, 1977, at Salisbury State College, Maryland. The first workshop "Models in Operations Research," will be led by Professor W. F. Lucas of Cornell. He was director of the MAA-NSF 1976 Faculty workshop on Modules in Applied Mathematics and taught NSF-AAAS Chautauqua-type short courses in 1975–76 and 1976–77. The second workshop, "Survey of Computer Science," will be led by Dr. W. J. Collins of SSC. His special interests are structured programming, artificial intelligence, and computing in the undergraduate curriculum. He was a guest lecturer at the 1976 MAA-SSC Summer Workshop and was rated by the participants as outstanding. These workshops are designed for teachers in two and four-year colleges. The total cost (including room and board) for each workshop is $120.00. For further information, write to Dr. B. A. Fusaro, SSC, Salisbury, Md. 21801 or call (301) 546–3261 Ext. 364.

DISCRETE OPTIMIZATION 1977

A sequence of 20 surveys given by leading experts of integer programming and covering its major directions will be presented in Vancouver (on the UBC campus) from August 8 to 12, 1977. Participation is open for mathematicians and practitioners of operations research interested in this area. Programme information is available from P. Hammer (University of Waterloo), E. Johnson (IBM, Yorktown Heights, New York), and B. Korte (University of Bonn). Because of the limited number of places, early registration is advisable with B. Alspach, D077 Department of Mathematics, Simon Fraser University, Burnaby, B.C., V5A 1S6, Canada. Campus housing is available. Combined charter flights for this symposium and the Seattle meeting of MAA-AMS may be possible.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

OFFICERS AND COMMITTEES AS OF FEBRUARY 1, 1977

General Offices: 1225 Connecticut Avenue, N. W., Washington, D. C. 20036
Executive Director: A. B. WILLCOX
Executive Director Emeritus: H. M. GEHMAN
Editorial Director: RAOUL HAILPERN
Secretary Emeritus: H. L. ALDER

OFFICERS

President, H. L. ALDER, University of California, Davis (1977–78)
Past-President, H. O. POLLAK, Bell Telephone Laboratories (1977–78)
First Vice-President, R. C. BUCK, University of Wisconsin (1976–77)
Second Vice-President,
Editor, R. P. BOAS, Northwestern University (1977–81)
Secretary, D. P. ROSELLE, VPI and State University (1975–79)
Treasurer, LEONARD GILLMAN, University of Texas (1973–77)

ADDITIONAL MEMBERS OF THE BOARD OF GOVERNORS

Ex-Presidents
R. P. BOAS, Northwestern University (1975–80)
VICTOR KLEE, University of Washington (1973–78)

Elected Members of the Finance Committee
R. D. ANDERSON, Louisiana State University (1974–77)
G. B. PRICE, University of Kansas (1976–79)

Governors-at-Large

ROY DUBISCH, University of Washington (1976–78)
DAVID GALE, University of California, Berkeley (1976–78)
CHARLES HATFIELD, University of Missouri at Rolla (1975–77)
MARY B. WILLIAMS, Ohio State University (1975–77)
(1977–79)
(1977–79)

Editor of the Mathematics Magazine

L. A. STEEN, St. Olaf College (1977)

Editor of the Two-Year College Mathematics Journal

C. A. LATHAN, Monroe Community College, Rochester, New York (1975–78)

Sectional Governors (July 1, 1974 — June 30, 1977)

Florida, R. C. MEACHAM, Eckerd College
Illinois, J. A. SCHUMAKER, Rockford College
Iowa, ELSIE C. MULLER, Morningside College
Louisiana-Mississippi, R. A. STOKES, University of Mississippi
Maryland-D.C.-Virginia, R. H. OWENS, University of Virginia
Michigan, R. V. CHURCHILL, University of Michigan
North Central, C. M. BRADEN, Macalester College
Philadelphia, DAVID ROSEN, Swarthmore College
Southern California, T. N. ROBERTSON, Occidental College
Texas, C. R. DEETER, Texas Christian University

Sectional Governors (July 1, 1975 — June 30, 1978)

Allegheny Mountain, E. F. MYERS, University of Pittsburgh
Indiana, M. J. MANSFIELD, Purdue University
Kentucky, D. B. COLEMAN, University of Kentucky
Metropolitan New York, W. L. ZLOT, New York University
Nebraska, MILDRED L. GROSS, Doane College
Northern California, G. L. ALEXANDERSON, University of Santa Clara
Oklahoma-Arkansas, J. M. JOBE, Oklahoma State University
Rocky Mountain, F. N. FISCH, University of Northern Colorado
Wisconsin, P. R. BENDER, Marquette University

Sectional Governors (July 1, 1976 — June 30, 1979)

Kansas, R. E. SHERMOEN, Washburn University
Missouri, C. J. STUTH, Stephens College
New Jersey, EILEEN L. POIANI, St. Peter's College
Northeastern, D. L. KREIDER, Dartmouth College
Ohio, R. L. WILSON, Ohio Wesleyan
Pacific Northwest, J. N. REAY, Western Washington State College
Seaway, ERIK HEMMINGSEN, Syracuse University
Southeastern, TREVOR EVANS, Emory University
Southwestern, J. E. NYMANN, University of Texas at El Paso

COMMITTEES OF THE ASSOCIATION

Terms of members expire, except where otherwise noted, at the Annual Meeting in January following the last year of service listed below. For temporary committees, no terms are listed since they are automatically discharged at the expiration of the President's term of office, which is at the Annual Meeting in January, 1978.

EXECUTIVE COMMITTEE

H. L. ALDER, *Chairman* (1977–78); R. P. BOAS (1977–81), R. C. BUCK (1976–77), LEONARD GILLMAN (1973–77), D. P. ROSELLE (1975–79), *Second Vice President*, (1977–78), all *ex officio*.

FINANCE COMMITTEE

H. L. ALDER, *Chairman* (1977–78), *ex officio*; R. D. ANDERSON (1974–77), LEONARD GILLMAN (1973–77), *ex officio*, H. O. POLLAK (1975–78), *ex officio*, G. B. PRICE (1976–79), D. P. ROSELLE (1975–79), *ex officio*.

Budget Review Subcommittee: R. D. ANDERSON (1974–77), G. B. PRICE (1976–79), both *ex officio*.

Investment Committee: LEONARD GILLMAN, *Chairman* (1973-77), *ex officio*; H. L. ALDER (1975-77), E. A. CAMERON (1976-78), C. V. NEWSOM (1976-78), D. P. ROSELLE (1975-79), *ex officio.*

COMMITTEE ON ADVISEMENT AND PERSONNEL

G. E. NOETHER, *Chairman* (1975-77); BERNICE L. AUSLANDER (1976-78), R. C. BUEKER (1976-78), CRAIG COMSTOCK (1976-78), JOAN P. LEITZEL (1977-79), C. J. NESBITT (1975-77), EILEEN L. POIANI (1975-77).

COMMITTEE ON CORPORATE MEMBERS

GORDON RAISBECK, *Chairman* (1975-77), D. R. BRILLINGER (1977-79), JANE K. CULLUM (1976-78), R. E. GASKELL (1975-77), LEONARD GILLMAN (1976-78), P. M. KAHN (1975-77), MARJORIE L. STEIN (1977-79), LEONARD TORNHEIM (1976-78), A. B. WILLCOX (1977-79).

COMMITTEE ON EARLE RAYMOND HEDRICK LECTURES

H. O. POLLAK, *Chairman* (1975-77); F. J. ALMGREN, Jr. (1976-78), M. D. DAVIS (1977-79).

COMMITTEE ON EDUCATIONAL MEDIA

D. I. SCHNEIDER, *Chairman* (1977-79); A. A. BLANK (1975-77), MIRIAM R. HECHT (1975-77), N. L. MAX (1975-77), *Director of film projects*; S. J. MILLES (1977-79), L. J. NACHMAN (1975-77), H. O. POLLAK (1975-77), SEYMOUR SCHUSTER (1977-79), J. E. SKELTON (1975-77).

COMMITTEE ON HIGH SCHOOL CONTESTS

R. A. ARTINO, *Chairman* (1977-79); WILLIAM DICE (1976-78), THOMAS ENGLERT (1977-79), A. M. GAGLIONE (1977-79), S. L. GREITZER (1975-77), M. S. KLAMKIN (1975-77), W. K. MCNABB (1976-78), N. S. MENDELSOHN (1975-77), W. E. MIENTKA, *Director*, (1977-79), R. S. PIETERS (1975-77), B. E. SCOTT (1977-79), NIEL SHELL (1975-77), J. R. SMART (1975-77), MARTHA ZELINKA (1977-79).

ADVISORY PANEL OF THE COMMITTEE ON HIGH SCHOOL CONTESTS

GEORGE BERZSENYI, E. A. CAMERON, N. A. CHILDRESS, A. L. DULMAGE, W. C. FOREMAN, HUBERT HUNZEKER, FRANK KOCHER, ABRAHAM SCHWARTZ, ARNOLD WENDT.

SUBCOMMITTEE ON THE USA MATHEMATICAL OLYMPIAD

S. L. GREITZER, *Chairman* (1977-79); R. A. ARTINO (1975-77), *ex officio*, GEORGE BERZSENYI (1977-79), TOM GRIFFITHS (1975-77), M. S. KLAMKIN (1977-79), W. K. MCNABB (1976-78), C. C. ROUSSEAU (1976-78), J. H. SPENCER (1976-78).

Advisor to the Olympiad Awards Ceremonies: NURA D. TURNER (1976-78).

COMMITTEE ON INSTITUTES AND WORKSHOPS

R. V. HOGG, *Chairman* (1976-78); W. F. LUCAS (1976-78), J. J. MALONE, Jr. (1975-77), M. W. POWNALL (1975-77), HELEN M. ROBERTS (1975-77), R. P. WALKER (1976-78), R. J. WEBER (1976-78).

COMMITTEE ON NATIONAL AWARDS AND PUBLIC REPRESENTATION

L. W. COHEN, *Chairman* (1976-78); H. L. ALDER (1977-78), *ex officio*, R. P. DILWORTH (1976-79), DEBORAH T. HAIMO (1977-79), MARK KAC (1975-77).

COMMITTEE ON NEW IDEAS

P. J. KNOPP, *Chairman* (1977-78); M. I. AISSEN (1976-78), WADE ELLIS (1976-78), NEWCOMB GREENLEAF (1976-78), T. N. ROBERTSON (1977-78), A. B. WILLCOX (1975-77).

COMMITTEE ON PUBLICATIONS

E. F. BECKENBACH, *Chairman* (1977-79); D. J. ALBERS (1976-78), F. J. ALMGREN, Jr. (1977-79), R. P. BOAS (1977-81), *ex officio*, ARGELIA V. ESQUIVEL (1975-77), D. T. FINKBEINER II (1975-77), LEONARD GILLMAN (1973-77), *ex officio*, ROBERT GILMER (1977-79), WANDA W. HELM (1975-77), C. A. LATHAN (1975-78), *ex officio*, IVAN NIVEN (1975-77), J. A. SEEBACH (1977), *ex officio*, A. C. TUCKER (1977-78).

SUBCOMMITTEE ON BASIC LIBRARY LISTS

D. FINKBEINER II, *Chairman*; R. P. BOAS, D. W. BUSHAW, H. H. HALL, R. H. MCDOWELL, G. B. PEDRICK, *ex officio.*

SUBCOMMITTEE ON CARUS MONOGRAPHS

D. T. FINKBEINER II, *Chairman* (1975-77), R. P. BOAS (1976-78), ARGELIA V. ESQUIVEL (1975-77), ROBERT GILMER (1977-79).

SUBCOMMITTEE ON DOLCIANI MATHEMATICAL EXPOSITIONS

R. A. HONSBERGER, *Chairman* (1975-77), D. J. ALBERS (1976-78), H. L. ALDER (1975-77).

SUBCOMMITTEE ON MAA STUDIES IN MATHEMATICS

G. L. WEISS, *Chairman* (1976–78); F. J. ALMGREN, Jr. (1977–79), WANDA W. HELM (1975–77), A. C. TUCKER (1975–77).

SUBCOMMITTEE ON MISCELLANEOUS PUBLICATIONS

E. F. BECKENBACH, *Chairman* (1977–79), LEONARD GILLMAN (1973–77), D. P. ROSELLE (1975–79), all *ex officio.*

SUBCOMMITTEE ON THE NEW MATHEMATICAL LIBRARY

IVAN NIVEN, *Chairman* (1976–77); W. G. CHINN (1977–79), BASIL GORDON (1977–79), M. M. SCHIFFER (1976–78).

SUBCOMMITTEE ON GEORGE POLYA AWARDS

JOSEPH HASHISAKI, *Chairman* (1977–79); E. F. BECKENBACH (1976–79), *ex officio,* BETTY J. HINMAN (1977–78).

SUBCOMMITEE ON LESTER R. FORD AWARDS

I. J. SCHOENBERG, *Chairman* (1975–77); E. F. BECKENBACH (1977–79), *ex officio,* B. L. ROTHSCHILD (1977–78).

SUBCOMMITTEE ON MATHEMATICS MAGAZINE AWARDS

ROY DUBISCH, *Chairman* (1977); E. F. BECKENBACH (1977–79), *ex officio,* N. J. A. SLOANE (1977–78).

COMMITTEE ON SECONDARY SCHOOL LECTURES

D. B. SMALL, *Chairman* (1976–78); G. L. ALEXANDERSON (1975–77), E. Z. ANDALAFTE (1977–79), D. M. HILL (1977–79), VIVIENNE M. MAYES (1977–79), DOROTHY P. SMITH (1976–78).

COMMITTEE ON SECTIONS

L. H. LANGE, *Chairman* (1977–80); J. C. BRADFORD (1977–80), L. A. GUILLOU (1976–79), S. W. HAHN (1975–78), JACQUELINE C. MOSS (1975–78), ARNOLD WENDT (1974–77), A. B. WILLCOX, *ex officio.*

COMMITTEE ON SPECIAL FUNDS OF THE ASSOCIATION

G. B. PRICE, *Chairman* (1975–77); E. F. BECKENBACH (1977–79), E. A. CAMERON (1976–78), H. M. GEHMAN (1975–77), LEONARD GILLMAN (1973–77), *ex officio,* W. L. HART (1976–78), B. W. JONES (1976–78), R. L. WILDER (1975–77), C. V. NEWSOM (1977–79).

COMMITTEE ON THE AWARD FOR DISTINGUISHED SERVICE TO MATHEMATICS

SAUNDERS MAC LANE, *Chairman* (1975–77); L. W. COHEN (1976–78), V. L. KLEE, Jr. (1977–79).

COMMITTEE ON THE CHAUVENET PRIZE

M. D. DAVIS, *Chairman* (1975–77); ERIK HEMMINGSEN (1977–79), L. A. ZALCMAN (1976–78).

COMMITTEE ON THE EXCHANGE OF INFORMATION ON MATHEMATICS

D. W. LICK, *Chairman* (1976–78); A. L. HAMMOND (1976–78), REUBEN HERSCH (1975–77), J. R. C. LEITZEL (1977–78), L. A. STEEN (1975–77).

COMMITTEE ON THE MEMBERSHIP OF THE ASSOCIATION

C. R. DEETER, *Chairman* (1977–79); H. L. ALDER (1977–78), *ex officio,* R. D. ANDERSON (1977–78), E. A. CAMERON (1977), LEONARD GILLMAN (1977), *ex officio,* P. A. LINDSTROM (1977–79), B. E. RHOADES (1977–78), D. P. ROSELLE (1977–79), *ex officio,* A. M. WEDEL (1977).

COMMITTEE ON THE PUTNAM PRIZE COMPETITION

J. D. E. KONHAUSER, *Chairman* (1975–77); R. T. BUMBY (1976–78), A. P. HILLMAN, *Director,* (1973–77), L. A. ZALCMAN (1977–79).

COMMITTEE ON THE UNDERGRADUATE PROGRAM IN MATHEMATICS

W. F. LUCAS, *Chairman* (1976–78); D. W. BUSHAW (1976–78), JUANITA J. PETERSON (1976–78), F. S. ROBERTS (1976–78), A. C. TUCKER (1976–78), G. S. YOUNG (1976–78).

COMMITTEE ON TWO-YEAR COLLEGES

D. J. ALBERS, *Chairman* (1977–79); W. G. CHINN (1975–77), J. E. CICERO (1977–79), E. P. COOPER (1976–78), R. M. DAVIS (1975–77), JOSEPH HASHISAKI (1975–77), BETTY J. HINMAN (1977–79), D. R. ILTIS (1975–77), C. A. LATHAN (1976–78), P. A. LINDSTROM (1976–78), JUNE P. WOOD (1975–77), H. E. ZINK (1976–78).

COMMITTEE ON VISITING LECTURERS AND CONSULTANTS

M. W. POWNALL, *Chairman* (1977–79); T. S. CHIHARA (1975–77), BRINDELL HORELICK (1977–79), MABEL D. MONTGOMERY (1975–77), W. M. MYERS (1976–78), MARY B. WILLIAMS (1975–77).

JOINT ADVISORY COMMITTEE ON GRADUATE PROGRAMS AT TRADITIONALLY BLACK INSTITUTIONS

C. B. BELL (AMS), LIPMAN BERS (AMS), R. C. BUCK (MAA), I. N. HERSTEIN (MAA), FRANK JAMES (NAM), TED SYKES (NAM).

JOINT COMMITTEE ON EMPLOYMENT OPPORTUNITIES

Terms of members of this committee expire on February 28 of the last year of service listed.

W. N. SMITH, *Chairman* (1976–80, MAA); R. D. ANDERSON (1976–80), E. C. POSNER (1976–78, SIAM), J. A. NOHEL (1974–78, AMS).

JOINT COMMITTEE ON PLACES OF MEETINGS

G. L. WALKER, *Chairman*; EVERETT PITCHER, D. P. ROSELLE, A. B. WILLCOX, all *ex officio*.

JOINT COMMITTEE ON THE EMPLOYMENT OF MATHEMATICIANS IN TWO-YEAR COLLEGES

J. M. SACHS, *Chairman*; LOUIS AUSLANDER, BETTY J. HINMAN, H. J. OSNER, E. H. SPANIER.

JOINT COMMITTEE ON THE TRAINING OF GRADUATE STUDENTS TO TEACH

J. H. WELLS, *Chairman*; M. W. HAM, NANCY T. HOLDER, E. E. MOISE, H. N. SHAPIRO, S. K. STEIN.

JOINT COMMITTEE ON WOMEN IN MATHEMATICS

AMS		MAA		SIAM	
ETTA Z. FALCONER	(79)	DOROTHY L. BERNSTEIN	(80)	JANE K. CULLUM	(79)
MARY W. GRAY	(77)	SHIRLEY A. HILL	(77)	CATHLEEN S. MORAWETZ	(78)
I. N. HERSTEIN	(78)	V. L. KLEE, Jr.	(79)		
ALICE T. SCHAFER, *Chairman*	(80)	JACQUELINE C. MOSS	(78)		

JOINT PROJECTS COMMITTEE IN MATHEMATICS

AMS	MAA	SIAM
SAUNDERS MAC LANE, *Chairman*	F. J. ALMGREN	G. F. CARRIER
G. D. MOSTOW	DOROTHY L. BERNSTEIN	H. G. COHEN
J. T. SCHWARTZ	W. H. KRUSKAL	P. R. GARABEDIAN

EDITORIAL BOARDS OF THE ASSOCIATION

AMERICAN MATHEMATICAL MONTHLY (all terms expire December 31, 1978)

Editor: R. P. BOAS (1977–81)

Associate Editors: JOSHUA BARLAZ, R. A. BRUALDI, D. Z. DJOKOVIĆ, R. K. GUY, RAOUL HAILPERN, SHIRLEY A. HILL, W. F. LUCAS, P. T. MIELKE, SEYMOUR SCHUSTER, J. A. SEEBACH, Jr., IVAR STAKGOLD, E. P. STARKE, L. A. STEEN, J. H. WELLS.

MATHEMATICS MAGAZINE (all terms expire December 31, 1980)

Co-Editors: J. A. SEEBACH, Jr., L. A. STEEN

Associate Editors: T. F. BANCHOFF, J. U. DUDLEY, D. J. EUSTICE, R. L. GRAHAM, RAOUL HAILPERN, R. A. HONSBERGER, R. E. HORTON, L. M. KELLY, MORRIS KLINE, P. J. MALRAISON, L. F. MEYERS, DORIS W. SCHATTSCHNEIDER.

TWO-YEAR COLLEGE MATHEMATICS JOURNAL (all terms expire December 31, 1978)

Editor: C. A. LATHAN

Board of Editors: DOROTHY L. BERNSTEIN, G. G. BITTER, J. C. DAVIS, R. C. FISHER, J. E. FORBES, S. A. GREENSPAN, RAOUL HAILPERN, JOSEPH HASHISAKI, BETTY J. HINMAN, ERWIN JUST, F. G. LASAK, P. A. LINDSTROM (NCTM representative), RALPH MANSFIELD, J. W. METTLER, NANCY MYERS (NCTM representative), JUNE P. WOOD.

REPRESENTATIVES OF THE ASSOCIATION

On Sections of the American Association for the Advancement of Science:

Section A: L. J. PAIGE (1975–77) Section Z: L. J. PAIGE (1975–77)

Section T: W. J. LeVeque (1975–77) Section U: G. E. Noether (1975–77)
Section X: A. B. Willcox (1975–77)

On the Advisory Committee to the Topology Films Project:

G. S. Young (1976–78)

On the Council of the Conference Board of the Mathematical Sciences:

H. O. Pollak, *ex officio*, G. B. Price, *ex officio*

On the Discipline Committee for Mathematics of the College Entrance Examination Board:

D. L. Kreider (1976–78)

On the Governing Council of Mu Alpha Theta:

R. L. Wilson (1976–78)

On the National Research Council:

Dorothy L. Bernstein (July 1, 1974 — June 30, 1977)

On the Steering Committee of the Calculus Project, Educational Development Corporation:

(1977–79)

On the U. S. Commission on Mathematical Instruction:

W. G. Chinn (July 1, 1973 — June 30, 1977), Mary P. Dolciani (July 1, 1974 — June 30, 1978).

OFFICERS OF THE SECTION

Allegheny Mountain

Chairman - J. B. Derr, West Virginia University
First Vice-Chairman -A. L. Rabenstein, Washington and Jefferson College
Second Vice-Chairman - F. T. Kocher, Jr., Pennsylvania State University
Secretary/Treasurer - J. W. Milsom, Butler County Community College

Florida

Chairman - A. W. Goodman, University of South Florida
Chairman-Elect - B. L. Brechner, University of Florida
Vice-Chairman - A. C. Shershin, Florida International University
Vice-Chairman - I. D. Bello, Hillsborough Community College
Secretary/Treasurer - F. L. Cleaver, University of South Florida

Illinois

Chairman - J. G. Christiano, Northern Illinois University
Chairman-Elect - J. D. Bradburn, Elgin Community College
First Vice-Chairman - Therese L. Butzen, William Rainy Harper College
Second Vice-Chairman - D. F. Clevidence, Carl Sandburg College
Secretary/Treasurer - H. C. Saar, Administrator, Plainfield Public School

Indiana

Chairman - M. C. Gemignani, Indiana University-Purdue University
Vice-Chairman - G. J. Sherman, Rose Hulman Institute of Technology
Secretary/Treasurer - D. E. Wilson, Wabash College

Intermountain

Chairman - E. A. Davis, University of Utah
First Vice-Chairman - S. K. Parker, Idaho State University
Second Vice-Chairman - Leslie C. Glaser, University of Utah
Secretary/Treasurer - D. R. Snow, Brigham Young University

IOWA

Chairman - J. L. CORNETTE, Iowa State University
Vice-Chairman - ELLEN OLIVER, Westmar College
Secretary/Treasurer - B. E. GILLAM, Drake University

KANSAS

Chairman - E. E. BEOUGHER, Ft. Hays Kansas State College
Vice-Chairman - F. S. BRENNEMAN, Tabor College
Secretary/Treasurer - ELLEN C. VEED, Ft. Hays Kansas State College

KENTUCKY

Chairman - H. G. ROBERTSON, Murray State University
Chairman-Elect - B. R. LANE, Eastern Kentucky University
Secretary/Treasurer - T. M. JENKINS, University of Louisville

LOUISIANA-MISSISSIPPI

Chairman - J. E. KEISLER, Louisiana State University
Mississippi Vice-Chairman - T. A. ATCHISON, Mississippi State University
Louisiana Vice-Chairman - C. R. SMITH, Northeast Louisiana University
Secretary/Treasurer - J. R. FOOTE, University of New Orleans

MARYLAND-D.C.-VIRGINIA

Chairman - R. M. DAVIS, Northern Virginia Community College
Chairman-Elect - O. M. THOMAS, U. S. Naval Academy
Vice-Chairman - J. M. SMITH, George Mason University
Vice-Chairman - RICHARD MOLLOY, U. S. Naval Academy
Secretary - R. C. DRAKE, Washington Technical Institute
Treasurer - HEWITT KENYON, George Washington University

METROPOLITAN NEW YORK

Chairman - GERALD FREILICH, Queens College, CUNY
Vice-Chairman for Four Year Colleges - MAURICE NADLER, Pace University
Vice-Chairman for Two Year Colleges - HELEN B. SINER, Staten Island Community College, CUNY
Vice-Chairman for High Schools - H. D. RUDERMAN, Hunter College High School
Secretary - LILY E. CHRIST, John Jay College of Criminal Justice, CUNY
Treasurer - HOWARD KLEIMAN, Queensborough Community College, CUNY

MICHIGAN

Chairman - E. A. TANIS, Hope College
Vice-Chairman - J. E. ADNEY, Jr., Michigan State College
Vice-Chairman - N. H. REMINGTON, Delta College
Secretary/Treasurer - DELIA KOO, Eastern Michigan University

MISSOURI

Chairman - F. W. WILKE, University of Missouri, St. Louis
Vice-Chairman - G. C. SCHRAG, Central Missouri State University
Secretary/Treasurer - H. K. STUMPFF, Central Missouri State University
Chairman High School Lecturers - E. Z. ANDALAFTE, University of Missouri, St. Louis
Committee Chairman for High School Contest - A. R. TINSLEY, Central Missouri State University
Past-Chairman - J. R. DOWNING, Southwest Missouri State University

NEBRASKA

Chairman - S. D. LUKE, Nebraska Wesleyan University
Past-Chairman - L. M. LARSEN, Kearney State College
Chairman-Elect - P. A. HAEDER, University of Nebraska, Omaha
Secretary/Treasurer - H. M. COX, University of Nebraska, Lincoln
Committee Chairman - LARRY STEPHENS, University of Nebraska, Omaha

NEW JERSEY

Chairman - B. M. KIERNAN, St. Peter's College
Vice-Chairman for Speakers - KENNETH KALMANSON, Montclair State College
Vice-Chairman for Two Year Colleges - H. W. GODDERZ, Union College
Vice-Chairman for Innovations - S. A. FOOTE, Rutgers University
Vice-Chairman for High-School Contest - S. L. GREITZER, Rutgers University
Secretary-Treasurer - M. E. WHITE, Stevens Institute of Technology
Associate Secretary-Treasurer - SUSAN G. MARCHAND, Kean College

NORTH CENTRAL

Chairman - G. E. BERGUM, South Dakota State University
Chairman-Elect - M. W. LEGG, Moorhead State University
Past-Chairman - K. F. CARLSON, St. Cloud State University
Secretary/Treasurer - L. A. GUILLOU, St. Mary's College

NORTHERN CALIFORNIA

Chairman - D. W. BARNETTE, University of California, Davis
Vice-Chairman - JANE M. DAY, College of Notre Dame
Secretary/Treasurer - N. H. FISHER, San Francisco State University
Program Chairman - K. R. REBMAN, California State University, Hayward

NORTHEASTERN

Chairman - G. P. MURPHY, University of Maine
Vice-Chairman - R. S. PIETERS, Groton School
Secretary/Treasurer - G. W. BEST, Phillips Academy

OHIO

Chairman - J. A. MURTHA, Marietta College
Chairman-Elect - W. H. BEYER, University of Akron
Secretary/Treasurer - GUS MAVRIGIAN, Youngstown State University
Program Chairman - MARION D. WETZEL, Denison University
Past-Chairman - R. G. LAATSCH, Miami University

OKLAHOMA-ARKANSAS

Chairman - T. H. FAY, Hendrix College
Past-Chairman - J. H. YATES, Central State University
First Vice-Chairman - V. M. SNOOK, Oral Roberts University
Second Vice-Chairman - LEON SEMKOFF, Claremore Junior College
Secretary/Treasurer - E. K. MCLACHLAN, Oklahoma State University

PACIFIC NORTHWEST

Chairman - H. E. ZINK, Lane Community College
Chairman-Elect - D. W. DETEMPLE, Washington State University
Vice-Chairman for Two Year Colleges - L. A. CURNUTT, Bellevue Community College
Vice-Chairman for Four Year Colleges - ROY DUBISCH, University of Washington
Secretary/Treasurer - J. O. HERZOG, Pacific Lutheran University

PHILADELPHIA

Chairman - E. A. KLOTZ, Swarthmore College
Vice-Chairman - DORIS W. SCHATTSCHNEIDER, Moravian College
Secretary/Treasurer - P. E. BEDIENT, Franklin and Marshall College

ROCKY MOUNTAIN

Chairman - D. D. BUSHNELL, Fort Lewis College
Vice-Chairman - V. A. NELSON, Metropolitan State College
Secretary/Treasurer - D. W. BALLEW, South Dakota School of Mines & Technology

SEAWAY

Chairman - F. D. PARKER, St. Lawrence University
First Vice-Chairman - P. T. SCHAEFER, SUNY, College at Geneseo
Second Vice-Chairman - R. A. MEILI, Mohawk Valley Community College
Secretary/Treasurer - E. C. STOPHER, SUNY at Oswego

SOUTHERN CALIFORNIA

Chairman - P. B. YALE, Pomona College
First Vice-Chairman - JOHN TODD, California Institute of Technology
Second Vice-Chairman - ALICE A. HUFFMAN, California State Polytechnic University
Program Chairman - R. F. BROWN, University of California, Los Angeles
Secretary/Treasurer - E. I. DEATON, San Diego State University

SOUTHEASTERN

Chairman - EMILIE V. HAYNSWORTH, Auburn University
Chairman-Elect - I. C. GENTRY, Wake Forest University
Vice-Chairman - R. A. DOBYNS, Carson-Newman College
Secretary/Treasurer - J. D. NEFF, Georgia Institute of Technology

SOUTHWESTERN

Chairman - A. H. SEEGLITZ, Phoenix College
Vice-Chairman - C. M. DUBBS, New Mexico Institute of Mining & Technology
Secretary/Treasurer - ALVIN SWIMMER, Arizona State University, Tempe

TEXAS

Chairman - G. R. BLAKLEY, Texas A & M University
First Vice-Chairman - H. L. ROLF, Baylor University
Second Vice-Chairman - R. G. DEAN, Stephen F. Austin State University
Secretary/Treasurer - J. C. BRADFORD, Abilene Christian College
Level I Director - B. D. LANGSTON, Tarrant County Junior College
Level II Director - A. D. BROCK, East Texas State University

WISCONSIN

Chairman - F. G. FLOREY, University of Wisconsin, Superior
Vice-Chairman - G. B. KLATT, University of Wisconsin, Whitewater
Secretary/Treasurer - J. T. RENFROW, Beloit College

ACADEMIC MEMBERS ELECTED INTO THE ASSOCIATION

In accordance with the amendment adopted at the Business Meeting of the Association at Stillwater on August 30, 1961, the Board of Governors, at its meeting on January 28, 1977, elected to membership the twenty-ninth set of applicants for academic membership (for election of the other twenty-eight sets, see the April, 1975 issue of the MONTHLY, page 448–9, the December, 1975 issue, page 1041, the April, 1976 issue, page 315, and the December, 1976 issue, page 840). Approval of election was given to the following applicants for academic membership:

Alma College
College of the Redwoods
University of Delaware
Hartwick College
Lane Community College
Manhattan College
University of Michigan-Flint
Northwest Nazarene College

Olivet College
Oregon College of Education
University of Saskatchewan
Wheaton College
Wilkes College
University of Wisconsin-Whitewater

DAVID P. ROSELLE, *Secretary*

OCTOBER MEETING OF THE NORTH CENTRAL SECTION

The 1976 fall meeting of the North Central Section of the MAA was held at South Dakota State University, Brookings, on October 22–23, 1976. There were 106 persons in attendance, including 62 MAA members.

The principal speaker was Leonard Gillman, University of Texas, Austin, Treasurer of the MAA. The title of his presentation was *Choosing a wife.* James Serrin, University of Minnesota, Minneapolis, was the invited speaker for the Friday evening session. He spoke on *Foundations of thermodynamics* or *WYAWTKA entropy BWATA.*

The following South Dakota State University faculty presided at sessions: Gerald Bergum, Chairman of the NCS/MAA, at the Saturday morning session and business meeting; Kenneth Yocum at the Friday evening session; and Larry Bennett at the Saturday afternoon session.

At the business meeting a progress report was given about preparation for the first North Central Section Summer Seminar. The seminar will be held at Bemidji State University, Bemidji, June 21–25, 1977.

The Saturday contributed papers included:

A short history of Zorn's Lemma, by Paul Campbell, Saint Olaf College, Northfield.

Tautologies in a fragment of the propositional calculus, by Lisl Gaal, University of Minnesota, Minneapolis.

Unique factorization rings with zero divisors, by Steve Galovich, Carleton College, Northfield.

A population model for Drosophila melanogaster, by Paul Froeschl, Saint Mary's College, Winona.

A form of data adjustment obtained from a solution of simultaneous linear equations, by Donald Searls, Office of Statistical Methods of the Education Commission of the States, Denver.

Life without associativity, by Pierre Malraison, Carleton College, Northfield.

A chessboard proof of Fermat's Two-Square Theorem, by Loren Larson, Saint Olaf College, Northfield.

A final comment on closing the loopholes, by Larry Bennett, South Dakota State University, Brookings.

During the Saturday afternoon session the 8-minute 16 mm color and sound film, "Newton's equal areas," by Bruce and Gatherine Cornwall, was shown, courtesy of Pierre Malraison and Carleton College.

LOUIS GUILLOU, *Secretary-Treasurer*

NOVEMBER MEETING OF THE SEAWAY SECTION

The Fall Meeting of the Seaway Section of the MAA was held at Broome Community College, Binghamton, N.Y., on November 6, 1976, with a registered attendance of 110 people, including 91 members of the Association.

Professor F. D. Parker, St. Lawrence University, Chairman of the Section, presided at the morning session and at one of the parallel afternoon sessions, with Professor Richard Meili, Mohawk Valley Community College, Second Vice-Chairman, presiding at the other afternoon session.

Dr. H. O. Pollak, Bell Laboratories, Murray Hill, New Jersey, President of the MAA, gave an invited lecture, "Relationship between the applications of mathematics and teaching of mathematics."

A workshop on "Constructive Geometry" was conducted by Eugene Mozier, LaSalle Institute, and Shirley Stanley, Schenectady County Community College.

Ten contributed papers were presented:

1. *Boolean algebra vs. linear algebra,* by J. E. Graver, Syracuse University.
2. *Verbing mathematical nouns: A case study of educational tensions,* by Larry Copes, Ithaca College.
3. *Some Cantor sets and Cantor functions,* by Judith Palagallo, Hartwick College.
4. *Octets of vectors and their applications,* by Mou-Ta Chen, State University College at Brockport.
5. *Convolution transforms on R^n*, by J. V. Peters, St. Bonaventure University.
6. *A theorem on Bernoulli numbers of order k,* by D. O. McKay, University of Western Ontario.
7. *Canonical reducible cubic forms in n ($\geqq 3$)-variables,* by Frank Servedio, McMaster University.
8. *Umbral methods and alternating series,* by A. P. Guinand, Trent University.
9. *π: Arc length or area?,* by J. F. Smith, LeMoyne College.
10. *Report and results of the workshop on the teaching of statistics,* by R. F. Barnes, State University College at Brockport.

EMMET STOPHER, *Secretary-Treasurer*

CALENDAR OF FUTURE MEETINGS

Fifty-seventh Summer Meeting, University of Washington, August 14–16, 1977.

Sixty-first Annual Meeting, Atlanta, Georgia, January 6–8, 1978.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN, St. Francis College, Loretto, Pennsylvania, April 22–23, 1977.

FLORIDA, early March. Deadline for paper titles 2 wks. bef. mtg.

ILLINOIS, Chicago Loop College, Chicago, May 6–7, 1977.

INDIANA, Wabash College, Crawfordsville, April 30, 1977.

INTERMOUNTAIN

IOWA, Drake University, Des Moines, April 22–23, 1977.

KANSAS, Tabor College, Hillsboro, April 2, 1977.

KENTUCKY, early April. Deadline for papers 6 wks. bef. mtg.

LOUISIANA-MISSISSIPPI, Friday-Saturday before February 20. Deadline for papers 3 mths. bef. mtg.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA, University of Maryland, College Park, April 30, 1977.

METROPOLITAN NEW YORK, Sarah Lawrence College, April 24, 1977.

MICHIGAN, Eastern Michigan University, Ypsilanti, May 6–7, 1977.

MISSOURI, University of Missouri, St. Louis, April 29–30, 1977.

NEBRASKA, Nebraska Wesleyan University, Lincoln, April 15–16, 1977.

NEW JERSEY, Union College, Cranford, April 30, 1977.

NORTH CENTRAL, North Hennepin Community College, Minneapolis, Minnesota, April 29–30, 1977.

NORTHEASTERN, Middlebury College, Middlebury, Vermont, June 1977.

NORTHERN CALIFORNIA, first or second Saturday in February.

OHIO, Denison University, Granville, April 15–16, 1977.

OKLAHOMA-ARKANSAS, Oral Roberts University, Tulsa, Oklahoma, April 1–2, 1977.

PACIFIC NORTHWEST, second Saturday in June. Deadline for papers 6 wks. bef. mtg.

PHILADELPHIA, Moravian College, Bethlehem, November 19, 1977.

ROCKY MOUNTAIN, Metropolitan State College, Denver, Colorado, April 29–30, 1977.

SEAWAY, State University College at Buffalo, May 6–7, 1977.

SOUTHEASTERN, University of Alabama, Huntsville, April 1–2, 1977.

SOUTHERN CALIFORNIA, first or second Saturday in March.

SOUTHWESTERN, Phoenix College, Phoenix, Arizona, April 22–23, 1977.

TEXAS, Baylor University, Waco, April 1–2, 1977.

WISCONSIN, University of Wisconsin, Oshkosh, April 29–30, 1977.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Washington, February 12–17, 1978.

AMERICAN MATHEMATICAL ASSOCIATION OF TWO YEAR COLLEGES

AMERICAN MATHEMATICAL SOCIETY, University of Washington, August 15–18, 1977.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION, University of North Dakota, Grand Forks, June 13–16, 1977.

ASSOCIATION FOR COMPUTING MACHINERY, Olympic Hotel, Seattle, Washington, October 17–19, 1977.

ASSOCIATION FOR SYMBOLIC LOGIC, Chicago Sheraton, Chicago, April 28–29, 1977.

ASSOCIATION FOR WOMEN IN MATHEMATICS

CANADIAN SOCIETY FOR HISTORY AND PHILOSOPHY OF MATHEMATICS/SOCIÉTÉ CANADIENNE D'HISTOIRE ET DE PHILOSOPHIE DES MATHÉMATIQUES, Hamilton, Ontario, June 2, 1977.

FIBONACCI ASSOCIATION

INSTITUTE OF MATHEMATICAL STATISTICS, Seattle, Washington, August 14–18, 1977.

MU ALPHA THETA

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, Cincinnati, Ohio, April 20–23, 1977.

OPERATIONS RESEARCH SOCIETY OF AMERICA, San Francisco Hilton, May 9–11, 1977.

PI MU EPSILON, University of Washington, August 14–16, 1977.

SCHOOL SCIENCE AND MATHEMATICS ASSOCIATION

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS

SPECIAL ANNOUNCEMENT

AMS-MAA-SIAM CONGRESSIONAL SCIENCE FELLOWSHIP

Applications are invited from candidates in the mathematical sciences for a Congressional Science Fellowship to be supported jointly by the American Mathematical Society, the Mathematical Association of America and the Society for Industrial and Applied Mathematics for the twelve-month period beginning 1 September 1977. The AMS-MAA-SIAM Fellow will serve, along with three or four Fellows selected by the American Association for the Advancement of Science and half a dozen Fellows sponsored by other scientific societies, under an annual program coordinated by the AAAS. The stipend for the Fellowship is $16,000, which may be supplemented by a small amount toward relocation and travel expenses. The overall AAAS program is described in the October-November 1976 CBMS *Newsletter*, page 56 (where, however, the stipend is incorrectly reported as $17,000 instead of $16,000) and also in the 7 January 1977 issue of *Science*, page 55.

As indicated there, Congressional Science Fellows spend their fellowship year working on the staff of an individual congressman or a congressional committee or in the congressional Office of Technology Assessment, the objective of the program being to enhance science-government interaction, the effective use of science in government, and the training of persons with scientific background for careers involving such use. Based on information on available congressional staff positions for Fellows gathered by the AAAS during the summer, each Fellow's assignment is worked out by him and the congressional office concerned following an intensive two-week orientation and interview procedure organized by the AAAS during which he or she encounters many facets of Congress, the Executive Branch, and people and organizations on the Washington scene. The AAAS provides advice and assistance during this process and remains in frequent and regular contact with all the Fellows throughout the fellowship year. More detailed information about the program as a whole may be obtained on request from Dr. Richard Scribner, Director, AAAS Congressional Science Fellow Program, 1776 Massachusetts Ave., N.W., Washington, D.C. 20036.

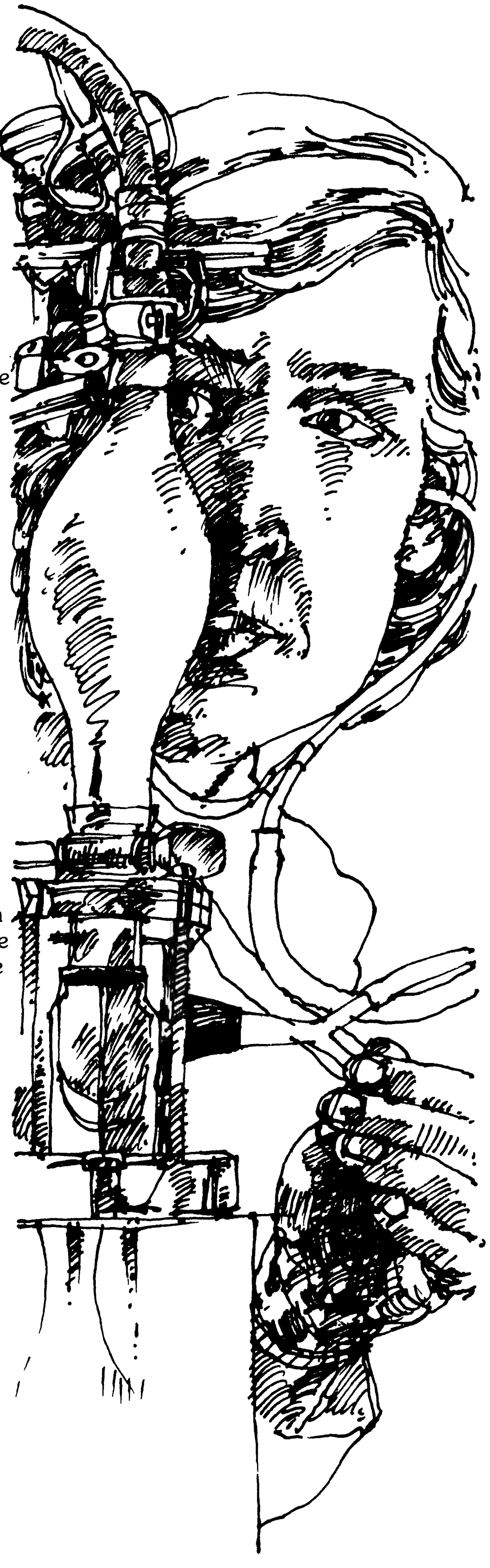
The AMS-MAA-SIAM Congressional Science Fellowship is to be awarded competitively to a mathematically trained person at the postdoctoral to mid-career level without regard to sex, race, or ethnic group. Selection will be made by a panel of the AMS-MAA-SIAM Joint Projects Committee for Mathematics, a nine-member committee consisting of three representatives from each of these organizations, with the cooperation and advice of Dr. Scribner. *Applications should be sent to the Conference Board of the Mathematical Sciences, 2100 Pennsylvania Ave., N.W., #832, Washington, D.C. 20037. The deadline for receipt of applications has been set at 21 May 1977*, and it is anticipated that the award will be made by late June.

In addition to demonstrating exceptional competence in some areas of the mathematical sciences, an applicant for the AMS-MAA-SIAM Congressional Science Fellowship should have a rather broad scientific and technical background and a strong interest in the uses of the mathematical and other sciences in the solution of societal problems. He or she should also be articulate, literate, flexible and able to work effectively with a wide variety of people. An application should state why the applicant wants to be a Congressional Science Fellow, should summarize his or her qualifications, and should be accompanied by a resume. Also, CBMS should receive by May 21 three letters from knowledgeable persons about the applicant's competence and suitability for the award.

Mathematics is for everyone

In your classroom today sit future surgeons, singers, sculptors, and perhaps even a scientist. Whatever the career, they all have one thing in common. They all need mathematics. At Addison-Wesley we publish quality texts at every level. Our authors are teachers who know what it takes to catch and hold student interest.

So whether their goal is a new face for Mount Rushmore or a cure for cancer — we can help them on their way.



Mathematics: Tools and Models

By Dalton R. Hunkins and Thomas L. Pirnot, *Kutztown State College*

An appealing alternative to the usual mathematics survey texts. Class-tested over several semesters, the authors found their modeling approach and realistic applications provided motivation and reinforcement for students. A unique three-part format for each chapter consists of a problem, tools of mathematics, and a solution; the object being the use of the tools to find the solution.

Calculus, Second Edition

By Lynn H. Loomis, *Harvard University*

This edition is a more traditional treatment of standard calculus topics, but it retains the flexibility and intuitive flavor of the first edition. Major changes include revision and simplification of topics such as inverse functions, the definite integral, and vectors; as well as relocation of the antiderivative chapter. Coverage of Green's Theorem has been added, and there is new material on the intermediate value principle.

Numerical Analysis

By Lee W. Johnson and R. Dean Riess, *Virginia Polytechnic Institute and State University*

A sophomore/junior level text for an introductory numerical analysis course. Features a broad range of topics including: solutions of overdetermined systems, optimization techniques, and the fast Fourier transform.

If you would like to be considered for complimentary examination copies or would like more information, write to Alfred Walters, Information Services, Addison-Wesley. Please include course title, enrollment, and author of text now in use.

Science & Mathematics Division

ADDISON-WESLEY PUBLISHING COMPANY

Reading, Massachusetts 01867

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

AN OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 84 NUMBER 5

CODEN: AMMYAE

CONTENTS

(*Continued on inside cover*)

MAY 1977

NOTICE TO AUTHORS

Research papers per se are unsuitable; see statement of policy (Vol. 81, p. 1).

Please follow the format in current issues of the MONTHLY. Manuscripts must be legibly typewritten or reproduced from typewritten copy, double spaced with wide margins and on one side of the paper. Three copies should be submitted to the appropriate editor and one kept by the author as protection against loss. The author's full address *must* appear at the end of the manuscript.

Backlog: Main Articles 19 months, Math. Notes 18 months, Research Problems 11 months, Classroom Notes 15 months, Math. Education 16 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to R. P. BOAS, Department of Mathematics, Northwestern University, Evanston, IL 60201; NOTES, etc.: to the corresponding Associate Editor; REPRINT PERMISSION: to LEONARD GILLMAN, Mathematical Association of America, University of Texas, Austin, Texas 78712 (see also the copyright notice below); ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N.Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N. W., Washington, D.C. 20036; BACK ISSUES: Contact P. and H. Bliss Co., Middletown, CT 06457.

Annual dues for members of the Association, including a subscription to the American Mathematical Monthly, are $18.00 for each of the first two years of membership and $22.00 thereafter. Student Membership is available with annual dues of $12.00. For nonmembers the subscription price is $25.00.

PUBLISHED BY THE ASSOCIATION at Washington, D.C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D.C., and additional mailing offices.

PRINTED IN THE UNITED STATES OF AMERICA

POSTER SESSIONS AT THE AUGUST 1977 MEETING OF THE ASSOCIATION

At its meeting at the University of Washington, August 14–16, 1977, the Association will sponsor a poster session in addition to the customary sessions of invited talks. The poster session is for contributed papers. Each contributor prepares a visual display. The displays are posted during an early part of the meeting and, at a later session, the author is present for discussions with interested persons.

Contributions from all areas of mathematics education at the collegiate level are invited for this poster session. Two-year college mathematicians are especially invited to contribute papers to this poster session.

Abstract forms and additional information are available upon request from David Roselle, Secretary, Mathematical Association of America, VPI&SU, Blacksburg, VA., 24061. The deadline for completed abstract is July 10, 1977.

MORLEY POLYGONS

FRANCIS P. CALLAHAN

Introduction: "*Pulchra dicuntur quae visa placent* — beauty is that which, being seen, pleases," [2]. This is Thomas Aquinas's definition of Beauty. It applies well to Mathematical Beauty in which lack of understanding (seeing) is so often responsible for lack of pleasure.

G. H. Hardy [3] gives two examples of readily accessible beautiful mathematical theorems. They are the theorem that the number of primes is infinite and the theorem that $\sqrt{2}$ is irrational. The list could be extended without great difficulty. Morley's Theorem (that adjacent angle trisectors of *any* triangle intersect in three points which are the vertices of an *equilateral* triangle) certainly belongs on the list. Classes and individuals to whom the author has introduced the theorem always react with some enthusiasm as soon as the statement of the theorem is understood. Coxeter [1] calls Morley's Theorem "one of the most surprising theorems of elementary geometry". Probably the main reason for the surprise and beauty of the theorem is the natural way in which it causes an *equilateral* triangle to emerge from *any* triangle.

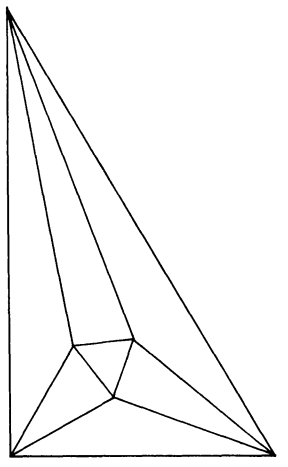


FIG. 0. Morley's Theorem

This paper generalizes Morley's Theorem. Define a Morley polygon to be a non-reentrant planar n-gon whose adjacent angle trisectors intersect to form the vertices of a regular n-gon. In these terms, Morley's Theorem asserts that every triangle is a Morley triangle. It is easy to see, also, that every regular polygon is a Morley polygon and that a rectangle not a square is not a Morley quadrilateral. The dotted line in Figure 1 shows a non-trivial Morley hexagon. This paper explores the possibilities and completely characterizes all Morley polygons.

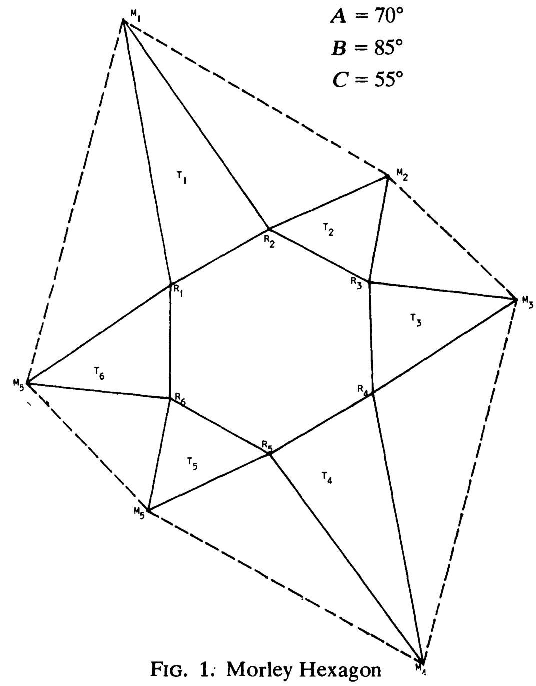


FIG. 1. Morley Hexagon
Ordinary nontrivial configuration
Generating angles (tooth angles for T_1)

Our proof is modeled to a considerable extent on that of Morley's Theorem as presented in [1]. The fundamental insight is that it is easier to start with the regular polygon and build the Morley polygon around it than to start with the Morley polygon and trisect its angles so as to arrive at the regular polygon.

First we show how to construct two types of Morley polygons around a given regular polygon. Then we show that the construction method is valid, and then we show that the Morley polygons so constructed are the only Morley polygons. Finally, we discuss briefly some extensions of these concepts.

1. The constructions. Two general types of Morley polygons exist, ordinary and snaggled. They are constructed as follows:

ORDINARY MORLEY POLYGONS. Let a regular n-gon R be given, n being divisible by 3. Define angle D by $D = \pi/n$, and select any three angles, A, B and C such that

$$A + B + C = \pi + D,$$

$$D < A < (2/3)\pi + D,$$

$$D < B < (2/3)\pi + D,$$

$$D < C < (2/3)\pi + D.$$

Then on the sides of the regular polygon R erect triangles external to R whose base angles are (A, B),

(B, C), (C, A), ... and so on in clockwise order with a periodicity of 3. Join the apices of the triangles in order and thus form a polygon, M. (Again, see Figure 1.) Then, as will be shown in the next section, M is a Morley polygon. Call it an *ordinary* Morley polygon.

When n is not divisible by 3 this construction fails for general values of A, B and C because the triangles do not "match up" when we travel round the regular polygon back to the starting point. In this case matching can only be achieved in the trivial case for which

$$A = B = C = (1/3)(\pi + D).$$

We also call this an ordinary Morley polygon, the only ordinary Morley polygon (except for irrelevant linear scale factor) for the case in which n is not divisible by 3. An ordinary Morley polygon for which $A = B = C$ we call a *trivial* one.

TERMINOLOGY AND NOTATION: Let the vertices of the regular polygon R be $R_1, R_2, \ldots, R_n$ with the proviso that R_1 and R_n may also be referred to as R_{n+1} and R_0, respectively, when convenient. Let the vertices of the polygon M be referred to as $M_1, M_2, \ldots, M_n$ with the same proviso about M_1 and M_n. Let T_i be the triangle with vertices R_i, R_{i+1}, and M_i. When M is a Morley polygon refer to the whole configuration, R, M, and $T_1, T_2, \ldots, T_n$ as a "Morley configuration."

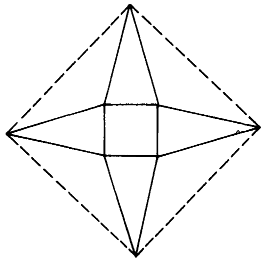
It is convenient to refer to the triangle T_i as the ith "tooth" of the configuration. Let the left and right base angles of T_i be designated by A_i and B_i and define the angle C_i by

$$C_i = \pi + D - A_i - B_i.$$

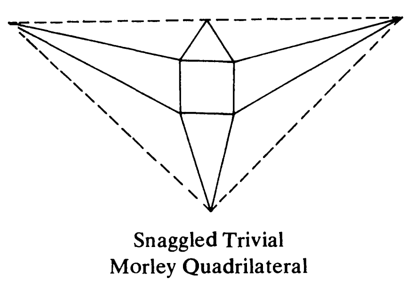
Then (A_i, B_i, C_i) will be called the "tooth angles" for the ith tooth. Note that C_i is *not* the apex angle of the triangle T_i; the apex angle is $C_i - D$.

SNAGGLED MORLEY POLYGONS: Other Morley polygons can be derived from the ordinary ones by a process we shall call "snaggling," as follows:

Let an ordinary Morley configuration be given and select a tooth, T_i, whose two base angles are acute and whose apex angle is less than $\pi/3$. That is, in terms of tooth angles, $A_i < \frac{1}{2}\pi$, $B_i < \frac{1}{2}\pi$, and $C_i - D < \pi/3$ so that $C_i < (\pi/3) + D$. Call these conditions the "snaggle conditions." It is not difficult

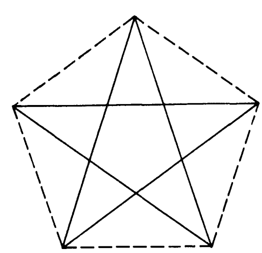


Ordinary Trivial
Morley Quadrilateral

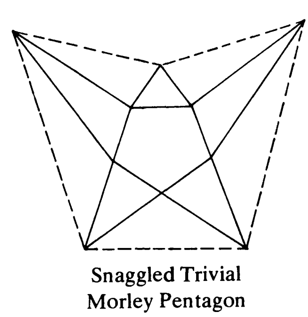


Snaggled Trivial
Morley Quadrilateral

FIG. 2



Ordinary Trivial
Morley Pentagon



Snaggled Trivial
Morley Pentagon

FIG. 3

to show that at least one tooth out of every three in an ordinary Morley configuration satisfies the snaggle conditions.

Now "snaggle" the set of three teeth, T_{i-1}, T_i, and T_{i+1}, by replacing them by teeth having base angles (B'_i, C'_i), (A'_i, B'_i) and (C'_i, A'_i), where A'_i, B'_i and C'_i, the tooth angles of the new tooth which replaces T_i are given by

$$A'_i = \tfrac{1}{2}\pi + D - B_i$$

$$B'_i = \tfrac{1}{2}\pi + D - A_i$$

$$C'_i = \pi - C_i.$$

(Some snaggled configurations are shown in Figures 2, 3, 4, and 5.) We append some data in the following self explanatory table:

		Generating Angles				Snaggle Angles	
n	D	A	B	C	A'	B'	C'
3	60°	80°	80°	80°	70°	70°	100°
4	45°	75°	75°	75°	60°	60°	105°
5	36°	72°	72°	72°	54°	54°	108°
6	30°	70°	70°	70°	50°	50°	110°
7	25.7°	68.6°	68.6°	68.6°	47.1°	47.1°	111.4°
8	22.5°	67.5°	67.5°	67.5°	45°	45°	112.5°

Data for some Snaggled Morley Configurations

As will be shown presently, the resulting configuration after snaggling is still a Morley configuration. The set of three teeth altered by snaggling will be called a "snaggle." As the figures show, an ordinary Morley configuration can be snaggled several times provided that the snaggles do not overlap.

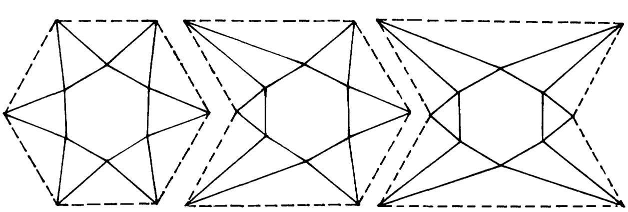
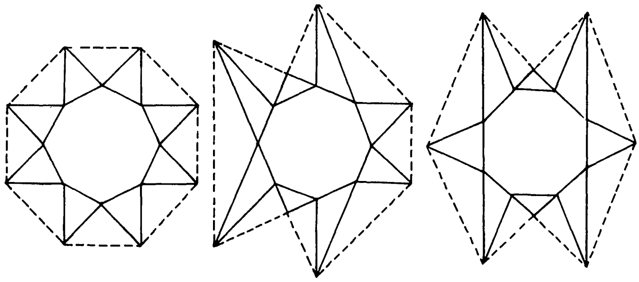


Fig. 4. Trivial Morley Hexagon, Ordinary and Snaggled

It may be worth observing that when n is divisible by 3 and all teeth are snaggled, the configuration again becomes ordinary. This is illustrated in Figure 4 in the twice snaggled hexagon. Also, when n is divisible by 3 and each snaggle is separated from its neighbor by a number of teeth divisible by 3, snaggled teeth can be viewed as unsnaggled and unsnaggled as snaggled. This is illustrated by Figure 4



Ordinary | Adjacent Snaggles Nonadjacent Snaggles

FIG. 5. Morley Octagon Configurations, Ordinary and with Two Snaggles

in which the once snaggled hexagon can be viewed as derived by snaggling from either the ordinary hexagon or from the twice snaggled hexagon.

2. Proof that the constructions work. In this section we show why the constructions work. We begin by considering the conditions that a single tooth of a Morley configuration must satisfy and then consider the conditions that a pair of adjacent teeth must satisfy.

SINGLE TOOTH CONDITIONS. The conditions that the tooth angles (A_i, B_i, C_i) of tooth T_i must satisfy are:

$$A_i + B_i + C_i = \pi + D$$

$$0 < A_i,\ 0 < B_i, \quad \text{and} \quad D < C_i < (2/3)\pi + D.$$

These conditions follow because A_i and B_i are the left and right base angles of the tooth and $C_i - D$ is the apex angle. This makes all but the last condition evident. The last condition is that $C_i - D$, the apex angle, should be less than $(2/3)\pi$. This must be true because this apex angle is one third of the ith vertex angle of the Morley Polygon M. This vertex angle must be less than 2π if the polygon is not to be reentrant.

DOUBLE TOOTH CONDITIONS. Now consider two adjacent teeth, (see Figure 6). In the figure, M_iR_i and $M_{i+1}R_{i+2}$ have been extended to meet at P. In the figure as drawn R_{i+1} is inside the triangle

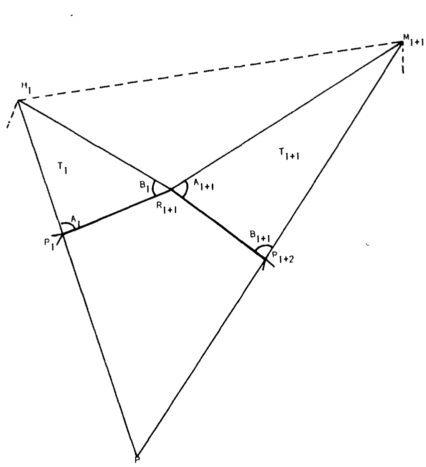


FIG. 6. Two Adjacent Teeth

$M_iM_{i+1}P$, but it could also happen that the two lines would be parallel or R_{i+1} might be outside the triangle. The discussion is similar in these cases and the conclusion is the same. We make no further reference to these other cases.

Since M_iR_i and M_iR_{i+1} are the trisectors of the ith apex angle of M when M is a Morley polygon, M_iR_{i+1} is the angle bisector of angle $R_iM_iM_{i+1}$; similarly, $M_{i+1}R_{i+1}$ is the angle bisector of angle $M_iM_{i+1}R_{i+2}$. (This line of argument is taken from the proof of Morley's Theorem found in **[1]** given at the end of the paper.) Conversely, if the analogous lines are angle bisectors for every adjacent pair of teeth then M is a Morley Polygon.

In the case shown in the figure the fact that R_{i+1} is the intersection of these angle bisectors makes R_{i+1} the incenter of triangle $M_iM_{i+1}P$. In **[1]** it is observed that the incenter O of a triangle ABC can be characterized as the unique point on the angle bisector of angle C such that the angle AOB is equal to $\frac{1}{2}\pi + \frac{1}{2}C$. Applying this observation to the triangle $M_iM_{i+1}P$ shows that R_{i+1} must be equidistant from sides M_iP and $M_{i+1}P$ and that

$$\text{angle } M_iR_{i+1}M_{i+1} = \tfrac{1}{2}\pi + \tfrac{1}{2}(\text{angle } R_iPR_{i+2}). \tag{1}$$

Since R is a regular n-gon length R_iR_{i+1} = length $R_{i+1}R_{i+2}$ and angle $R_iR_{i+1}R_{i+2} = \pi - 2D$, where, as before, $D = \pi/n$. Letting the common length of a side of R be r we have

$$r \sin A_i = r \sin B_{i+1}$$

when R_{i+1} is equidistant from the sides of the triangle. That is,

$$\text{Either E: } B_{i+1} = A_i \quad \text{or} \quad \text{S: } B_{i+1} = \pi - A_i. \tag{2}$$

As indicated, we call these the E-case (equal) and the S-case (supplementary), respectively.

Next, by considering the quadrilateral $R_iR_{i+1}R_{i+2}P$ it is easy to show that

$$\text{angle } R_iPR_{i+2} = A_i + B_{i+1} + 2D - \pi.$$

Using these relationships in (1) and rearranging gives

$$A_{i+1} = \pi + D - \tfrac{1}{2}(A_i + B_{i+1}) - B_i. \tag{3}$$

Using (2) in this equation gives for the relations between the base angles of adjacent teeth in the E-case and the S-case,

$$\begin{array}{lll} & \text{E-}\mathit{case} & \text{S-}\mathit{case} \\ (4) & A_{i+1} = \pi + D - B_i - A_i & A_{i+1} = \tfrac{1}{2}\pi + D - B_i \\ & B_{i+1} = A_i & B_{i+1} = \pi - A_i \end{array}$$

By making use of the single-tooth conditions,

$$A_i + B_i + C_i = A_{i+1} + B_{i+1} + C_{i+1} = \pi + D,$$

we can rewrite (4) as follows:

$$\begin{array}{lll} & \text{E-}\mathit{case} & \text{S-}\mathit{case} \\ & A_{i+1} = C_i & A_{i+1} = \tfrac{1}{2}\pi + D - B_i \\ (5) & B_{i+1} = A_i & B_{i+1} = \pi - A_i \\ & C_{i+1} = B_i & C_{i+1} = \tfrac{1}{2}\pi + D - C_i \end{array}$$

The formulas (5) will be useful shortly.

A vertex R_{i+1} between tooth T_i and tooth T_{i+1} will be called an E-vertex if the tooth angles of T_i and T_{i+1} are related by the E-formulas above and an S-vertex if they are related by the S-formulas.

Note that when $A_i = \frac{1}{2}\pi$ both sets of formulas become the same. In this case we call the vertex an E-vertex.

We summarize the discussion in the following:

THEOREM 1: *A configuration of the type being considered is a Morley configuration if and only if*
(1) *At each tooth the tooth angles satisfy*

$$A_i + B_i + C_i = \pi + D \text{ and}$$
$$0 < A_i,\ 0 < B_i,\ D < C_i < (2/3)\pi + D.$$

(2) *Each vertex of R is either an* E*-vertex or an* S*-vertex.*

It now follows easily that the ordinary and snaggled configurations described in Section 1 are indeed Morley configurations. For the ordinary configuration all the vertices of R are E-vertices (this is particularly clear from formulas (5) above) and the conditions on generating angles A, B and C insure that the single-tooth conditions are met at every vertex. Snaggled configurations consist of sequences of snaggled teeth separated by sequences of unsnaggled teeth. An R-vertex between two teeth of the same type (both snaggled or both unsnaggled) is an E-vertex and an R-vertex between two teeth of different type is an S-vertex. Example: For a snaggle centered at T_i and not adjacent to any other snaggle the sequence of vertices R_{i-1}, R_i, R_{i+1}, and R_{i+2} is SEES. For two adjacent snaggles centered at T_i and T_{i+3} and not adjacent to any other snaggles the sequence of vertices $R_{i-1}, R_i, \ldots, R_{i+5}$ is SEEEEES. Finally, the snaggle conditions insure that teeth T_{i-1}, T_i, and T_{i+1} will satisfy the single-tooth conditions after snaggling.

In summary:

COROLLARY. *Ordinary and snaggled configurations are indeed Morley configurations.*

3. Proof that there are no other Morley polygons. In this section we prove that there are no Morley polygons other than the ordinary and snaggled ones considered already.

E–S *sequences.* We shall call the sequence of E's and S's attached to the vertices of the R of a Morley configuration starting with R_2 the E–S sequence for the configuration. Because of the formulas (5) it is clear that a knowledge of the tooth angles for T_1 together with a knowledge of the E–S sequence for the configuration is enough to determine the configuration completely. We shall abbreviate an E–S sequence by writing E^k and S^k when k consecutive E's or S's appear in the sequence. This is appropriate because E and S can be viewed as operators that transform the angles of one tooth into those of the next.

An ordinary Morley configuration, then, is one whose E–S sequence consists entirely of E's. A snaggled Morley configuration is one such that

(1) There is at least one tooth whose angles would also serve as generators of an ordinary Morley configuration.

(2) If we take that tooth to be T_1, the E–S sequence is of the form

$$\mathrm{E}^{j_1}(\mathrm{S}\mathrm{E}^{i_1}\mathrm{S})\mathrm{E}^{j_2}(\mathrm{S}\mathrm{E}^{i_2}\mathrm{S})\mathrm{E}^{j_3}(\mathrm{S}\mathrm{E}^{i_3}\mathrm{S})\mathrm{E}^{j_4}\ldots,$$

where each i has remainder 2 when divided by 3 and each j is 0 or a positive integer.

Conversely, as is easy to see, a Morley configuration that satisfies this description can always be obtained by snaggling an ordinary Morley configuration. Thus, to show that there are no Morley configurations other than ordinary and snaggled, it suffices to show that any Morley configuration whose E–S sequence contains at least one S is of this type. Or if we call a Morley configuration which contains at least one S in its E–S sequence an extraordinary Morley configuration then we must show that every extraordinary Morley configuration is snaggled.

The z-plane. In order to demonstrate this result it is convenient first to change notation slightly and then to associate a complex number, z_i, with the ith tooth. The use of complex numbers is not essential to the argument but it does serve to simplify some formulas.

First the notation change: Let a_i, b_i, c_i, and d be the complements of A_i, B_i, C_i, and D, resp., so that

$$a_i = \tfrac{1}{2}\pi - A_i,\ b_i = \tfrac{1}{2}\pi - B_i,\ c_i = \tfrac{1}{2}\pi - C_i, \quad \text{and} \quad d = \tfrac{1}{2}\pi - D.$$

In terms of these angles the single-tooth conditions and the E and S formulas become:

$$a_i + b_i + c_i = d,\ a_i < \tfrac{1}{2}\pi,\ b_i < \tfrac{1}{2}\pi, \quad \text{and} \quad d - (2/3)\pi < c_i < d$$

and

$$\text{(6)} \qquad \begin{array}{ll} \text{E-}\textit{case} & \text{S-}\textit{case} \\ a_{i+1} = c_i & a_{i+1} = d - b_i = a_i + c_i \\ b_{i+1} = a_i & b_{i+1} = -a_i = -a_i \\ c_{i+1} = b_i & c_{i+1} = d - c_i = a_i + b_i \end{array}$$

The second set of formulas for the S-case is of some slight interest because they show that both E and S transformations can be expressed homogeneously without reference to d or π.

Next define complex number z_i by the equation

$$z_i = (a_i\omega + b_i\omega^2 + c_i)/d, \tag{7}$$

where ω is the cube root of unity: $\omega = -\frac{1}{2} + \frac{1}{2}\sqrt{3}i$ ($i = \sqrt{-1}$ here).

Given z_i, the values of a_i, b_i and c_i can be calculated as follows:

Write

$$dz_i = a_i\omega + b_i\omega^2 + c_i$$
$$d\bar{z}_i = a_i\omega^2 + b_i\omega + c_i$$
$$d = a_i + b_i + c_i$$

($\bar{z}$ is the complex conjugate of z). These equations are easily solved to yield

$$a_i = \frac{d}{3}[2\mathrm{Re}(z_i\omega^2) + 1]$$

$$b_i = \frac{d}{3}[2\mathrm{Re}(z_i\omega) + 1]$$

$$c_i = \frac{d}{3}[2\mathrm{Re}(z_i) + 1] \qquad \text{(Re means "real part of")}.$$

Use of these formulas in the single-tooth conditions shows that for z to correspond to a tooth of a Morley configuration we must have:

$$\mathrm{Re}(z\omega^2) < 1 + \frac{3}{n-2}$$

$$\mathrm{Re}(z\omega) < 1 + \frac{3}{n-2}$$

$$-\left(1 + \frac{4}{n-2}\right) < \mathrm{Re}(z) < 1.$$

(We have also used: $d = \frac{1}{2}\pi - \pi/n$.)

It is not difficult to use (6) to calculate the effect of an E or S transformation on z:
If R_{i+1} is an E-vertex, then $z_{i+1} = \omega z_i$.
If R_{i+1} is an S-vertex, then $z_{i+1} = -\bar{z}_i - \omega^2$.
Put another way, if we treat E and S as operators,

$$z\mathrm{E} = \omega z$$

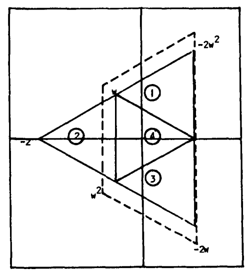
$$z\mathrm{S} = -\bar{z} - \omega^2.$$

Note that we have written the operator to the *right* of the operand z, as is sometimes done. This avoids some confusion in the present situation.

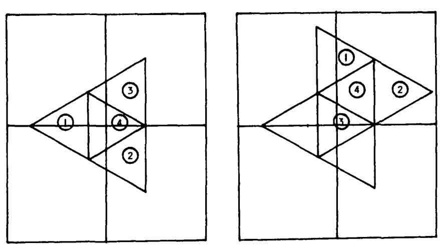
The operator E rotates z counterclockwise through an angle of $(2/3)\pi$. To describe S graphically we write:

$$(zS - \tfrac{1}{4}) = -(\bar{z} - \tfrac{1}{4}) + \tfrac{1}{2}\sqrt{3}i.$$

From this it is easy to see that S reflects z in the vertical line $x = \frac{1}{4}$ and translates it upward by amount $\frac{1}{2}\sqrt{3}$.



The z-plane
Dotted line shows admissible region.



Motion caused by E Motion caused by S

FIG. 7. Effect of E and S

To visualize further the effect of E and S on the z-plane refer to Figure 7. Note that S causes the image of cell 4 to move to the position previously occupied by cell 1.

One conclusion about E–S sequences is now evident; since the repeated action of S moves z steadily upward, S^k can never return z to its original position (for $k \neq 0$); since z_{n+1} must be the same as z_1 in a Morley configuration, *an* E–S *sequence can never consist entirely of* S's.

To investigate further we consider the motion of points in the z-plane when various sequences of E's and S's are applied. When a point moves outside the admissible region (defined by the single-tooth conditions for z) we consider it to be annihilated and note that a sequence that eventually annihilates all points in the z-plane is impossible because then no z could be taken for z_1.

To simplify the discussion we replace each sequence before beginning by a "reduced sequence" gotten by using $E^3 = 1$ (1 is the identity transformation) so as to allow no higher power of E than E^2 to appear in the sequence; further we ignore all constraints except $\mathrm{Re}(z) < 1$ in defining the admissible

region. The effect of using the reduced sequence is to consider a more fragmentary history of z than would be obtained by using the full sequence. For example, if the sequence were S^2E^5S, the reduced sequence would be S^2E^2S and instead of considering $z_1, z_2, \ldots, z_8$ we would only consider z_1, z_2, z_3, z_4, z_5. The effect, thus, of these simplifications is to relax the conditions to be satisfied by an E–S sequence. Nevertheless, they are sufficient to establish the desired result.

The exact result now to be established is:

THEOREM 2. *Let a reduced* E–S *sequence of a Morley configuration operate on* z, *to produce* $z_2, z_3, \ldots, z_m$. *If the sequence contains at least one* E *and one* S *and if* $\mathrm{Re}(z_i) < 1$ *for* $i = 1, 2, \ldots, m$, *then the sequence is of the form*

$$(SE^2S)E^{j_1}(SE^2S)E^{j_2}(SE^2S)E^{j_3}\ldots,$$

where each j *is either* 0, 1 *or* 2. *Further, (possibly after a cyclic permutation of the sequence)* z_1 *can be taken to satisfy* $\mathrm{Re}(\omega^k z_1) < 1$ *for* $k = 0, 1, 2$. *Finally, if the number of members of the reduced sequence is* m, *then* $z_1E^m = z_1$.

This theorem implies the desired result that every extraordinary Morley configuration is a snaggled Morley configuration. As an aid to seeing this, note that the condition $\mathrm{Re}(\omega^k z_1) < 1$ for $k = 0, 1, 2$ puts z_1 in cell 4 in Figure 7, and that for a z in cell 4, z, zE and zE^2 all satisfy the single-tooth condition; also since $z_1E^m = z_1$, so also, $z_1E^n = z_1$, n being the length of the unreduced E–S sequence. This shows that Z_1 generates an ordinary Morley configuration. The remainder of the details in showing that the theorem implies the desired result are easy.

Now we prove the theorem in several steps:

(1) SES^kES $(k > 0)$ *cannot appear in an admissible reduced sequence.*

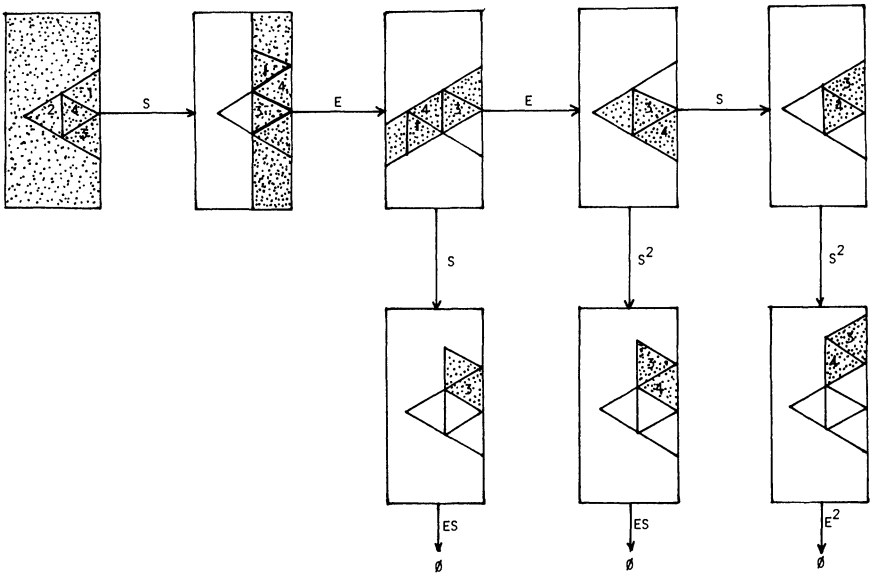


FIG. 8. Proof that SES^kES, $SES^{k+1}E^2S$, $SE^2S^{k+1}ES$, and $SE^2S^{k+2}E^2S$ are impossible for $k > 0$

Proof. See Figure 8. There the progress of points in the z-plane is illustrated. The right edge of each rectangle in the figure is the line $x = 1$ and points that pass to the right are considered to be annihilated. Thus in the first rectangle all points are considered admissible (so that the whole rectangle is shaded). After S moves the points, only the shaded vertical strip of the second rectangle contains

admissible points. Thus, as the figure shows, application of SESES leads to the annihilation of all the points which are originally admissible. ($\varnothing$ represents the null set.) For $k > 1$ the figure is similar.

(2) $SE^2S^{k+1}ES$ *and* $SES^{k+1}E^2S$ *are possible for admissible reduced sequences only when* $k = 0$.

Proof: Again see Figure 8. This gives the proof for $k = 1$ for the first sequence above; the proof for $k > 1$ is similar. The truth of the assertion for the second sequence follows because it is evident from symmetry that the reverse of an admissible sequence is admissible.

(3) $SE^2S^{k+2}E^2S$ *is admissible only when* $k = 0$.

The proof is similar.

From these deductions it follows that S can never appear to a power higher than the second; and of course neither can E in a reduced sequence. Thus, the most general reduced sequence can be written as

$S^{k_1}E^{m_1}S^{k_2}E^{m_2}\ldots$ where each k and m can have values 1 or 2.

Thus the following cases arise:

Case 1: $m_i = 1$ for at least one m in the sequence. In this case, because of (2) above, SE^2SESE^2S must appear in the sequence. (If the sequence is not long enough to allow this go around the polygon more than once.)

Case 2: $k_i = 2$ for at least one k in the sequence.

(Note that a sequence may belong to both case 1 and case 2).

In this case, because of (3) above, $SE^2S^2E^2S$ must appear somewhere in the sequence.

Case 3: Neither case 1 nor case 2. This is, $m_i = 2$ and $k_i = 1$ for every i. The sequence must therefore be $SE^2SE^2SE^2\ldots$.

Now a glance at the top row in Figure 8 shows that the effect of SE^2S is to take cell 4 to itself, to move cell 3 to the position formerly occupied by cell 1 and to annihilate the remainder of the z-plane. Thus, after writing the sequences in cases 1 and 2 as $(SE^2S)E(SE^2S)$ and $(SE^2S)^2$ we see that only cell 4 remains (and in its original position, though not necessarily with its original orientation) when either sequence is applied. In either case next in the sequence must come E^kS where $k = 0$, 1 or 2. Application of this operator moves cell 4 to the position initially occupied by cell 1. If anything is to survive, E^2S, which moves cell 4 back to its initial position, must come next. This done, the argument can be repeated. Thus, in cases 1 and 2, the sequence form is

$$(SE^2S)E^{k_1}(SE^2S)E^{k_2}\ldots \text{ with each } k = 0, 1, \text{or } 2.$$

Further the initial z, i.e., z_1, can be taken to be in cell 4.

The sequence in case 3 is already in the desired form (with $k_i = 2$ for all i). Now, however, cells 3 and 4 are brought back to their initial positions by the application of $(SE^2S)E^2$, so that it is possible that z_1 might be in cell 3 or in cell 4. If z_1 is in cell 3 then z_4, which is z_1SE^2 is in cell 4 and we can start the sequence using z_4 instead of the original z_1.

Thus, in all three cases the sequence has the desired form and, without loss of generality, z_1 can be taken to be in cell 4.

To complete the proof of the theorem we have only to show that $z_1E^m = z_1$, where m is the length of the reduced sequence. To show this we make use of the easily proven identity (for z_1 in cell 4) $SE^2S = E^4$ to replace SE^2S every time it appears in the original sequence before it is reduced. This changes the original sequence to a pure-E or ordinary sequence, from which it follows that $z_1E^n = z_1$. Since m differs from n by a multiple of 3 and since $E^3 = 1$ the result now follows.

With this the theorem is proven and, except for a few remarks, the paper is complete.

4. Remarks on the triangle. It might be well to show explicitly how Morley's original theorem follows from the discussion given here. Let A, B and C be the generating angles for an ordinary Morley configuration ($n = 3$), so that the vertex angles, M_1, M_2, and M_3 are $3(C - D)$, $3(B - D)$ and

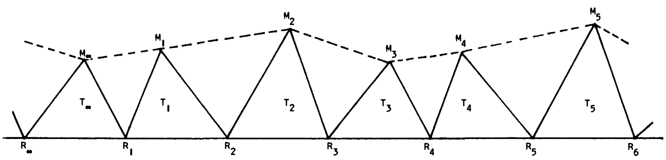
$3(A-D)$ where $D=(1/3)\pi$. Letting these angles be A'', B'', and C'' and solving for A, B, and C gives

$$A = A''/3 + D, \quad B = B''/3 + D, \quad \text{and} \quad C = C''/3 + D.$$

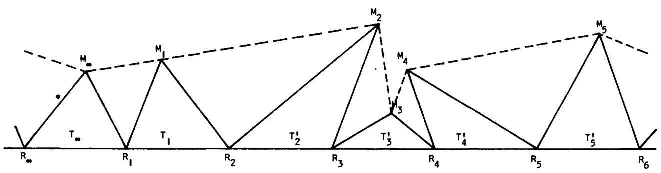
Since $0 < A'', B'', C'' < \pi$, this shows that $D < A, B, C < (2/3)\pi + D$. Since $A'' + B'' + C'' = \pi = 3D$, we have $A + B + C = \pi + D$. Thus, generating angles can be found to produce any triangle. That is, every triangle is similar to, and thus is, a Morley triangle.

Another remark on the triangle: If we allow the E–S sequence to go around the triangle twice, then $(SE^2S)E^2$ or its equivalent, $(SE^2)^2$, appears as a possible sequence for the triangle. However, in this case, z_1 would have to satisfy $z_1SE^2 = z_1$ (since z_4 and z_1 are identical for $n=3$). That is, $\omega^2(-\bar{z}_1 - \omega^2) = z_1$, or, after dividing by ω and transposing, $2\mathrm{Re}(z_1\omega^2) = -1$. This gives (recall the formulas for a, b and c,) $a_1 = 0$ and $A_1 = \frac{1}{2}\pi$. However, as was remarked already, E and S are indistinguishable when $A_1 = \frac{1}{2}\pi$ and S might just as well be called E. Thus, the case SE^2 does not really arise for the triangle and there are, thus, no extraordinary triangles.

5. The infinite case. As n approaches infinity, D approaches 0, and the Morley configuration becomes a sequence of triangles arranged along a straight line. Snaggling is still possible, with the



Ordinary nontrivial Morley Polygon for $n = \infty$



Snaggled nontrivial Morley Polygon for $n = \infty$

Generating angles (tooth angles for T_1)	$A = 70°$ $B = 50°$ $C = 60°$	Snaggle angles (Tooth angles for T_3')	$A' = 30°$ $B' = 40°$ $C_4' = 110°$

FIG. 9

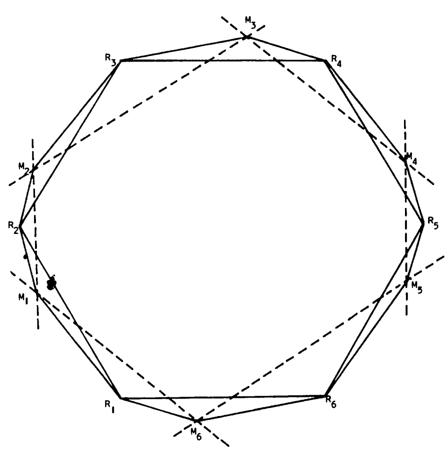


FIG. 10. Exterior Morley Hexagon (Type 1)

exception that the trivial case (all teeth become equilateral triangles) can no longer be snaggled. An example is shown in Figure 9.

6. Further work. Two students at the Penn State Graduate Center have discovered two different types of related polygons which they have called exterior Morley polygons.

The first of these, due to Joseph W. Owsley [4] is an n-gon in which the points of intersection of the trisectors of adjacent exterior angles meet to form a regular n-gon. In this case, an exterior angle is defined to be the angle between a side and the extension of an adjacent side. (See Figure 10 for an exterior Morley hexagon of this type.)

The second type, due to Catherine M. Myers [5] is defined in the same words, except that in this case the exterior angle is the external angle between two adjacent sides. (See Figure 11 for an ordinary non-trivial exterior Morley nonagon of this type and Figure 12 for a snaggled version.)

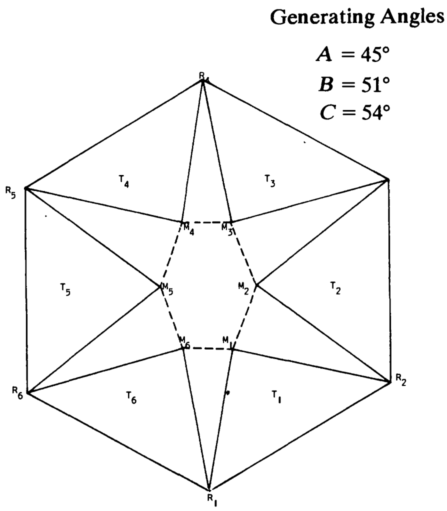


FIG. 11. Ordinary Nontrivial Exterior Morley Hexagon (Type 11)

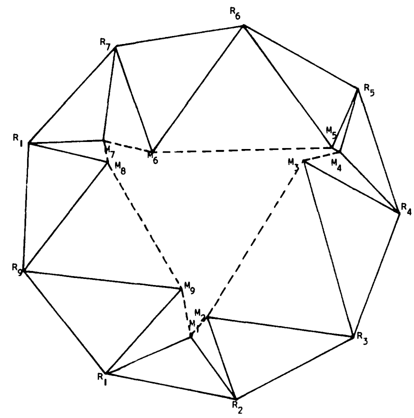


FIG. 12. Snaggled Exterior Morley Nonagon (Type 11)

Acknowledgements. Problem E1030 in this MONTHLY reawakened the author's interest in Morley's Theorem and was thus the immediate occasion for this investigation. A preliminary version of this paper was sent to the MONTHLY in connection with a comment on the problem. Professor S. P. Lipshitz, of the University of Waterloo, Ontario, wrote me that there was an error in the paper in that the possibility of angles A_i and B_{i+1} in Figure 6 of being supplementary had been overlooked. He enclosed an example due to Martin and Judah Milgram in which the angles were in fact supplementary. This initiated the study of the snaggled Morley polygons reported in this paper. The author is deeply grateful to Professor Lipshitz for pointing out this error in the preliminary version and for passing on to him the example due to the Milgrams.

References

1. H. S. M. Coxeter, Introduction to geometry, Wiley, New York, 1961.
2. Thomas Aquinas, Summa Theologica 1a, 5, 4, ad 1. (Cited in Copelston, A history of philosophy, Image Books, 1962.)
3. G. H. Hardy, A mathematician's apology, Cambridge Univ. Press, New York, 1967.
4. J. W. Owsley, Exterior Morley polygons, Penn State Master's Thesis, August, 1975.
5. C. M. Myers, Exterior Morley polygons, Penn State Master's Thesis, November, 1975.

KING OF PRUSSIA GRADUATE CENTER, PENNSYLVANIA STATE UNIVERSITY, KING OF PRUSSIA, PA 19406.

KLEINIAN TRANSFORMATION GEOMETRY

RICHARD S. MILLMAN

In this article we shall give a modern interpretation of transformation geometry. This subject has recently become of great interest to mathematics educators for use in kindergarten to high school, but has been paid too little attention at the college level. The usual approach to transformation geometry [**5**] or [**12**] consists of giving the classical geometries and then presenting their transformation groups. This certainly runs contrary to the ideas of the founder of transformation geometry, Felix Klein (1849–1929), who believed very strongly in a unified approach to mathematics whenever possible. I shall adopt the view that the proper approach to transformation geometry is through the geometry of homogeneous spaces and that this affords a modern interpretation of Klein's program (the Erlanger Programm) as outlined in his original paper of 1872 [**7**]. Roughly speaking, Klein's program says that a geometry on a space determines a group of transformations of the space and that a group of transformations on the space determines a geometry. While most modern mathematicians have been using the first idea (that very valuable information can be gained by looking at the group of transformations which leave the geometry invariant) we have neglected the second one. I claim that the theory of homogeneous spaces affords us a modern interpretation of the second point of Klein's program. I am not claiming that Klein foresaw the theory of homogeneous spaces but rather that with the help of the works of Riemann and E. Cartan (1869–1951) we can make precise (via the theory of homogeneous spaces) the notion that an arbitrarily chosen group will determine a geometry.

If we are to adopt Klein's approach to transformations, what definition are we to take for geometry? There is certainly no easy answer to this question but the approach of G. F. B. Riemann (1826–1866) is certainly the most modern in spirit and is the one I shall use here. Riemann's inaugural address [**11**] begins:

> "As is well known, geometry presupposes the concept of space, as well as assuming the basic principles for constructions in space. It gives only nominal definitions of these things, while their essential specifications appear in the form of axioms. The relationship between these presuppositions [the concept of space, and the basic properties of space] is left in the dark; we do not see whether, or to what extent, any connection between them is necessary, or *a priori* whether any connection between them is even possible."

This means that we must first decide what "space" should be. In section 2 we define space to be a homogeneous space; that is, the quotient space of a topological group G by a subgroup L so that $M = G/L$. In section 3 we assume that G is a subgroup of the group of nonsingular matrices (that is, G and L are Lie groups). In this section we add some geometric structure to that of space, which we call a geometry on the homogeneous space, and so obtain the notion of "lines." Throughout the last two sections of the paper we stress the fact that this definition of geometry includes in it, as special cases, Euclidean, Spherical and Hyperbolic geometries. We treat these three special cases in detail. (I do not claim, however, that this is the most general definition of geometry.) The first section of this paper gives a brief historical background before studying homogeneous spaces and their geometry.

Where do homogeneous spaces belong in the realm of mathematics? They are not just used as an interpretation of Klein's Erlanger program; they are also used to do function theory (harmonic analysis [**6**]), to serve as models in differential geometry [**8**], and are used in mathematical physics [**2**]. It is as Klein remarks in the notes he added (in 1893) to the original Erlangen address ([**7**], p. 244):

> "A model, whether constructed and observed or only vividly imagined, is for this geometry not a means to an end, but the subject itself."

I would like to thank George Parker for valuable conversations about this paper.

1. Klein's Erlanger Programm. In order to understand the importance of Klein's *Erlanger Programm* one must understand the state of geometry in the middle of the nineteenth century. There were on the one hand the synthetic geometers like Poncelet, Steiner and von Staudt, and on the other

hand analytic geometers like Plücker, Cayley and Grassmann. There was projective, descriptive and inversive geometry. In addition to these, there was the ancient question of the independence of the parallel postulate of Euclid. Lobachevsky and J. Bolyai (in about 1829) found geometries in which the fifth postulate was violated. Geometry was indeed very fragmented. The only successful attempt at unification in geometry was Riemann's brilliant inaugural address of 1854 [**11**], parts of which were quoted in the introduction to this paper. Most mathematicians of the nineteenth century found Riemann's lecture much too abstract and difficult to follow and so ignored it. For an excellent survey of nineteenth century geometry see [**3**, Chapter 24]. For a more complete account see [**4**]. For a modern interpretation of Riemann's paper see [**10**].

Felix Klein was very disturbed by the fragmentary state of geometry as can be seen from the following passages from the Erlangen address:

> "But it has seemed the more justifiable to publish connective observations of this kind, because geometry, which is after all one in substance, has been only too much broken up in the course of its recent rapid development into a series of almost distinct theories, which are advancing in comparative independence of each other... (p. 216)
>
> "When we consider, for instance, how persistently the mathematical physicist disregards the advantages afforded him in many cases by only a moderate cultivation of the projective view, and how, on the other hand, the student of projective geometry leaves untouched the rich mine of mathematical truths brought to light by the theory of the curvature of surfaces, we must regard the present state of mathematical knowledge as exceedingly incomplete and, it is to be hoped, as transitory." (p. 244)

One of Klein's fortes was the ability to penetrate enormous amounts of mathematics and extract the important points. When Klein was appointed professor at Erlangen in 1872, he used this ability to deal with the problem of geometry. Klein had recently visited Paris where he learned of the importance of group theory. He was also (collaborating with the Norwegian Sophus Lie (1842–1899)) using group theory to attack problems in differential equations. It was clear to Klein that the use of groups in geometry would unify and simplify much of geometry much as the use of groups in algebra helped clarify and unify the theory of equations (e.g., Galois theory) and his work with Lie helped to clarify some questions in differential equations. Klein's program had very little circulation at first. It was translated into both English (1893, [**7**]) and Italian (*Annali di Matematica*, **17** (1889)) and gained much wider circulation. This coupled with Klein's move to the University of Göttingen in 1886 and his reknown as a teacher as well as a researcher lent tremendous force to his Erlanger program and to the influence it was to have.

As we mentioned briefly in the introduction, there are two main points to Klein's *Erlanger Programm*. The first (which is advice we all follow today) is that any structure which is studied should be studied through the group of automorphisms which preserve the structure. Klein saw that a geometry could (and should) be studied by studying the group of transformations or motions which preserve the geometry (which he called the "principal group"). In his own words [**7**] (Klein uses "sense" for "orientation" and "manifoldness" for "space" or "manifold." The italics are his.):

> "Now there are space-transformations by which the geometric properties of configurations in space remain entirely unchanged. For geometric properties are, from their very idea, independent of the position occupied in space by the configuration in question, of its absolute magnitude, and finally of the sense in which its parts are arranged. The properties of a configuration remain therefore unchanged by any motions of space, by transformation into similar configurations, by transformation into symmetrical configurations with regard to a plane (reflection), as well as by any combination of these transformations. The totality of all these transformations we designate as the *principal group* of space-transformations; *geometric properties are not changed by the transformations of the principal group.* And, conversely, *geometric properties are characterized by their remaining invariant under the transformations of the principal group.*

From the above it is clear that the geometry determines the principal group. Klein's second point (which is all too often ignored) is that the converse is also true. A group which acts on a space determines a geometry. In his words:

> "*Given a manifoldness and a group of transformations of the same; to investigate the configurations belonging to the manifoldness with regard to such properties as are not altered by the transformations of the group.*" (p. 218)

It is this second point which we shall discuss in this paper. We shall see that to give a space M and a group G acting transitively on M is really the same as giving G and a subgroup L of G (which is the isotopy subgroup of G) such that $M = G/L$ (see Theorem 2.6). Furthermore, we see that M comes equipped with a group of transformations. (These will be called T_g for $g \in G$ in the next section.) We will define a notion of geometry on G/L in such a way that "lines" on G/L are carried to lines by T_g for each $g \in G$. In defining "geometry" on G/L we shall make use of some ideas from the work of E. Cartan (connections on homogeneous spaces). We will therefore see that a "manifoldness and a group of transformations of the same" do determine a geometry.

This approach is a unifying one, for as Klein remarks:

> "Particular stress is to be laid upon the fact that the choice of the group of transformations to be adjoined is quite arbitrary, and that consequently all the methods of treatment satisfying our general condition are in this sense of equal value." (p. 219)

It should be pointed out that this approach is not the most general one. The approach of Riemann **[11]** that uses a manifold and a quadratic differential for a line element is much more general. One difference between the two approaches is that most Riemannian manifolds have a discrete group of transformations (they are called "isometries" in the Riemannian case), whereas all of our spaces have very large (continuous) groups of transformations! (In the Riemannian case one uses the term "geodesic" instead of "line.") Our approach through homogeneous spaces is, however, general enough to include the classical geometries and so unifies most of the geometry of the nineteenth century. Klein, in fact, does discuss in [7] how the geometries of the nineteenth century are subsumed in his program. We shall examine, as examples, Euclidean, Spherical and Hyperbolic geometries in detail at the end of this paper.

2. Space in the Klein Problem. A group G is a *topological group* if G is also a topological space, such that both $\mu: G \times G \to G$ and $\iota: G \to G$ given by $\mu(g_1, g_2) = g_1g_2$ and $\iota(g) = g^{-1}$ are continuous. We shall now give some examples of topological groups all of which are matrix groups. We use $M(n)$ for the set of all $n \times n$ matrices.

We define the topology on a matrix group as follows: If $A \in M(n)$ then

$$A = \begin{pmatrix} a_{11} \cdots a_{1n} \\ \vdots \qquad \vdots \\ a_{n1} \cdots a_{nn} \end{pmatrix}$$

However, we may string out the elements of A and rewrite A as $(a_{11}, \ldots, a_{1n}, a_{21}, \ldots, a_{2n}, \ldots, a_{n1}, \ldots, a_{nn})$ which belongs to R^{n^2}. Define the metric on $M(n)$ to be the same as the metric on R^{n^2}. (This means that $d(A, B) = \sqrt{\sum_{j=1}^n \sum_{i=1}^n (a_{ij} - b_{ij})^2}$, where A and B belong to $M(n)$. Thus, a matrix group has a topology as a subset of R^{n^2}. It is easy to see that with this topology $Gl(n)$, $Sl(n)$ and $SO(n)$ (see Table 1) are topological groups since inversion and multiplication of matrices are polynomials in the entries. $A(n)$ may be viewed as a subset of $Gl(n+1)$ (and so has a topology) as follows: $A \in A(n)$ means that there are $B \in Gl(n)$ and $\boldsymbol{b} \in R^n$ such that for $\vec{v} \in R^n$, $A\vec{v} = B\vec{v} + \vec{b}$. Writing $B = (b_{ij})$ and $\vec{b} = (b_1, \ldots, b_n)$ we may identify $v \in R^n$ with $(v, 1) \in R^{n+1}$. $A \in A(n)$ is then identified with the element of $Gl(n+1)$

$$\begin{pmatrix} b_{11} \cdots b_{1n} & b_1 \\ b_{21} \cdots b_{2n} & b_2 \\ \vdots \qquad \vdots & \vdots \\ b_{n1} \cdots b_{nn} & b_n \\ 0 \ \cdots 0 & 1 \end{pmatrix}$$

TABLE 1: Some Common Matrix Groups

Group, G	Description (with $A \in G$)	Geometric Meaning
$Gl(n)$ (General linear group)	$n \times n$ matrices with $\det A \neq 0$	(a) Non-singular matr. (b) Sends basis to basis
$Sl(n)$ (Special linear group)	$n \times n$ matrices with $\det A = 1$	(a) Volume Preserving (b) $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ $ad - bc = 1 \mid (n = 2)$
SO(n) (Special orthogonal group)	$n \times n$ matrices with $AA^T = I$ (i.e. $A^T = A^{-1}$) and $\det A = 1$	(a) Rotations of R^n which leave origin fixed and preserve orientation (b) Sends orthonormal basis to o.n. basis (c) $(n = 2)$ $A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ with $a^2 + b^2 = 1$
$A(n)$ (Affine group)	$A: R^n \to R^n$ which are given by: $A\vec{v} = B\vec{v} + \vec{b}$ where $B \in Gl(n), \vec{b} \in R^n$	Rotations, Dilations and Translations of R^n

The collection of all such matrices has a topology as a subset of $Gl(n+1) \subset M(n+1)$. If U and V are subsets of G then we will use UV for the subset of G given by $\{uv \mid u \in U, v \in V\}$.

LEMMA 2.1. *If U is an open subset of G and V is any subset of G then both VU and UV are open in G.*

Proof: We define left multiplication by $g \in G$, $T_g: G \to G$, via $T_g(x) = gx$ where $x \in G$. If $i_g: G \to G \times G$ is given by $i_g(x) = (g, x)$ then i_g is continuous for all $g \in G$, hence $T_g = \mu \circ i_g$ is continuous for each $g \in G$. Because $(T_g)^{-1} = T_{g^{-1}}$ we see that T_g is a homeomorphism; in particular T_g must be an open map. Therefore if U is open in G then $T_g(U) = gU$ is open in G. Because $VU = \bigcup_{g \in V} gU$, VU is also open. The proof for UV is similar. ■

Contained in the above proof is the following:

COROLLARY 2.2. *If $T_g: G \to G$ is defined by $T_g(x) = gx$ then T_g is a homeomorphism of G for all $g \in G$.*

A subgroup L of G is always assumed to have the subspace (relative) topology. With this topology L is a topological group. Let $X = G/L = \{gL \mid g \in G\}$ be the usual coset space and let $\pi: G \to G/L$ be the canonical projection $\pi(g) = gL$ if $g \in G$. A subset $\mathcal{O}$ of G/L is defined to be open if $\pi^{-1}(\mathcal{O})$ is open in G. $X = G/L$ with this topology (which is called the quotient topology) is called the *homogeneous space defined by* G and L. We shall use the above notation throughout this paper.

PROPOSITION 2.3. (a) *π is a continuous, open map.* (b) *Let Y be a topological space and $\psi: G \to Y$ be continuous. If $\bar{\psi}: G/L \to Y$ is a set map such that $\psi = \bar{\psi} \circ \pi$ then $\bar{\psi}$ is continuous.* (c) *X is Hausdorff if and only if L is closed in G.*

Proof: (a) follows directly from Lemma 2.1. (b) follows from part (a). For a proof of (c) see [**9**, p. 180]. ■

Let $\bar{e} = eL \in X$ and $\bar{T}_g: X \to X$ be given by $\bar{T}_g(aL) = gaL$ for $a \in G$.

PROPOSITION 2.4. *$\bar{T}_g$ is a homeomorphism of X for each $g \in G$.*

The proof is a straightforward verification that $\bar{T}_g$ is well-defined since $(\bar{T}_g)^{-1} = \bar{T}_{g^{-1}}$. We shall now drop the "bar" from the notation. If $l \in L$ note that $T_l(gL)$ need not be gL although $T_l(\bar{e}) = \bar{e}$.

Philosophically, we have now defined "space," X, and "motions," T_g, of the space in the sense of Klein. Note that in this interpretation space comes equipped with motions but there is no notion of lines yet. We must add something in order to get the idea of lines. This is in keeping with the ideas of Riemann's paper [**11**]. Before putting lines on space we shall present three examples. The concept that unifies these examples is that of a topological transformation group.

Let G be a topological group and X a topological space. If there is a continuous map from $G \times X \to X$ (which we write $(g, x) \to (g \cdot x)$) such that for all $x \in X$ (1) $(gh) \cdot x = g \cdot (h \cdot x)$ for all $g, h \in G$ and (2) $e \cdot x = x$ then G is called a *topological transformation group* on X, or more simply, we say that G *acts* on X. G acts *transitively* on X if there is a point $x_0 \in X$ such that for each $y \in X$ there is a $g \in G$ with $g \cdot x_0 = y$. If G acts transitively on X then given any $x, y \in X$ there is $g \in G$ such that $g \cdot x = y$.

Example 1: Let $S^n = \{(a_1, \ldots, a_{n+1}) \in R^{n+1} \mid a_1^2 + \cdots + a_{n+1}^2 = 1\}$ be the n-sphere and let $e = (1, 0, \ldots, 0) \in S^n$. If $A \in SO(n+1)$ and $x \in S^n$ then $A \cdot x = Ax$ is to be interpreted as multiplying the unit $(n+1)$-vector x by the matrix A. Since A preserves the length of a vector, $Ax \in S^n$. Geometrically this action is transitive because $SO(n+1)$ is the set of rotations of R^{n+1}. Clearly we can rotate any given vector $x \in S^n$ to any other vector $y \in S^n$. More algebraically, if $x, y \in S^n$ complete to any orthonormal bases $X = \{x, x_1, \ldots, x_n\}$ and $Y = \{y, y_1, \ldots, y_n\}$ of R^{n+1}. Define $A: R^{n+1} \to R^{n+1}$ by $Ax = y$ and $Ax_i = y_i$, $i = 1, \ldots, n$. A is certainly orthogonal. If $\det A = -1$ then interchange the last two elements of Y and do the same thing.

Example 2: Let $H^2 = \{(x, y) \in R^2 \mid y > 0\} = \{x + iy \mid y > 0\}$ be the upper half plane. If $A \in Sl(2)$ then

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

with $ad - bc = 1$. If $z \in H^2$ then $A \cdot z$ is defined by $Az = (az + b)/(cz + d)$. (This is motivated by regarding $Sl(2)$ as a subgroup of the group of Möbius of Fractional Linear Transformations as in complex variables [**1**].) If $z = x + iy$ then the imaginary part of $A \cdot z$ is $y/[(cx + d)^2 + (cy)^2]$ which is positive if $y > 0$ so that $A \cdot z \in H^2$ if $z \in H^2$.

LEMMA 2.5. *$Sl(2)$ acts transitively on H^2.*

Proof: Let $e + if \in H^2$ so that $f > 0$. We show that there is an $A \in Sl(2)$ such that $A \cdot i = e + if$. Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \sqrt{f} & e\sqrt{f}/f \\ 0 & \sqrt{f}/f \end{pmatrix}$$

A belongs to $Sl(2, R)$. An easy computation shows that $Ai = e + if$. ■

Example 3: $A(n)$ acts on R^n by the definition of the Affine group. The action is transitive (take $x_0 = \mathbf{0}$, given $y \in R^n$ consider A defined by $A\vec{v} = \vec{v} + \vec{y}$ for $\vec{v} \in R^n$. A is what is classically called a pure translation.)

If G acts on X then the *isotropy subgroup of* G at $x_0 \in X$ is $L = \{g \in G \mid g \cdot x_0 = x_0\}$. L is a subgroup of G which is closed if X is Hausdorff. In example 1, with $x_0 = e$, we see that

$$L = \left\{A \in SO(n+1) \mid A = \begin{pmatrix} 1 & 0 \cdots 0 \\ 0 & \\ \vdots & \boxed{\tilde{A}} \\ 0 & \end{pmatrix} \text{ where } \tilde{A} \in SO(n)\right\}.$$

Note that L is isomorphic to $SO(n)$. In example 2, with $x_0 = i$, we see that

$$L = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \Big| \, a^2 + b^2 = 1 \right\}.$$

Note that L is isomorphic to $SO(2)$. In example 3 with $x_0 = \mathbf{0}$, $L = \{\mathbf{v} \to B\mathbf{v} + \mathbf{b} \mid \mathbf{b} = 0, B \in Gl(n)\}$. Note that L is isomorphic to $Gl(n)$.

Now assume that G acts transitively on X, $x_0 \in X$ and L is the isotropy subgroup of G at x_0. Consider $\psi: G \to X$ given by $\psi(g) = g \cdot x_0$. ψ induces a map $\bar{\psi}: G/L \to X$ given by $\bar{\psi}(gL) = \psi(g) = g \cdot x_0$. $\bar{\psi}$ is well-defined because if $g_1L = g_2L$ then $g_2^{-1}g_1 \in L$ so $g_2^{-1}g_1 \cdot x_0 = x_0$ or $g_1 \cdot x_0 = g_2 \cdot x_0$. $\bar{\psi}$ is continuous by Proposition 2.3b. Note that $g_1 \cdot x_0 = g_2 \cdot x_0$ if and only if $g_2^{-1}g_1 \in L$ so $\bar{\psi}$ is one to one. $\bar{\psi}$ is onto because G acts transitively on X. $\bar{\psi}: G/L \to X$ is therefore a continuous bijection. It is also open if mild assumptions are put on the topology of G and X. The precise theorem (whose proof requires the Baire Category Theorem, see [**9**, p. 185]) is:

THEOREM 2.6. *If G is a second countable, locally compact topological group and X is a locally compact topological space and G acts transitively on X then $\bar{\psi}: G/L \to X$ is a homeomorphism where L is the isotropy subgroup of a point.*

Since all three of our examples satisfy the hypotheses of the theorem we have three examples of homogeneous spaces (see Table 2). Actually in example 1 we may eschew the technical difficulties involved in proving that $\bar{\psi}$ is open. Since $SO(n+1)$ is a closed and bounded subspace of $R^{(n+1)^2}$, $SO(n+1)$ and hence $SO(n+1)/SO(n)$ are compact. $\bar{\psi}$ is now a continuous map from a compact space onto a Hausdorff space and hence is closed. Because $\bar{\psi}$ is also one to one it is a homeomorphism. *Ad hoc* proofs that $\bar{\psi}$ is open can also be presented for examples 2 and 3. Note that these three examples are the spaces which serve as models for the three classical geometries (spherical, hyperbolic, and Euclidean). There are, of course, other models for these geometries, in particular a frequently used model for hyperbolic geometry is the disk $D^2 = \{(x, y) \mid x^2 + y^2 < 1\}$. We prefer however to use H^2. In the table which follows, $\bar{A}$ means the coset represented by $A \in G$.

TABLE 2: Classical Homogeneous Spaces

Coset Space	Realization $\bar{\psi}$	Classical Name
$S^2 = SO(3)/SO(2)$	$\bar{A} \rightsquigarrow A\begin{pmatrix}1\\0\\0\end{pmatrix} \in S^2$	Riemann Sphere
$H^2 = Sl(2)/SO(2)$	$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ $\bar{A} \rightsquigarrow \dfrac{ai+b}{ci+d} \in H^2$	Hyperbolic (Poincaré) Upper Half Plane
$R^2 = A(2)/Gl(2)$	$\bar{A} \rightsquigarrow A(\vec{0}) \in R^2$	Euclidean Plane

3. Lines in space — geometry. In keeping with the Riemann program we must add to our "space" the concept of "line." In keeping with the Klein program we must insure that each of the transformations $T_g: X \to X$ carry lines to lines. Whatever lines are, our intuition says that through each point of X and in each "direction" there should be a unique line. It will suffice to define lines at $\bar{e} = eL \in X$ because we shall define a line at gL to be the image of a line at $\bar{e}$ under T_g. This means that we need only define direction at $\bar{e}$. As a start we define directions at $e \in G$. We shall assume that G is a subgroup of $Gl(n)$ for some n. We shall give a more precise condition later. Note that $e = I$, the identity matrix and that $G \subset R^{n^2}$ so it makes sense to differentiate any G-valued curve. (See the

example below.) In defining direction, we are motivated by the fact that in R^2 a direction is a vector; that is, a direction is the tangent to some curve in R^2. Let J be any interval about $0 \in R$. If $A: J \to G$ is differentiable and $A(0) = I$ then a *direction at I of G* is $X = A'(0)$.

As an example, consider

$$A(t) = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \in SO(2).$$

$$A'(t) = \begin{pmatrix} -\sin t & \cos t \\ -\cos t & -\sin t \end{pmatrix} \text{ so that } X = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

is a direction at I of $SO(2)$.

PROPOSITION 3.1. *If $\hat{G}$ is the set of all directions at I of G then* (a) *$\hat{G}$ is a finite dimensional real vector space,* (b) *if L is a subgroup of G then $\hat{L}$ is a vector subspace of $\hat{G}$.*

Proof: Let $A(t)$, $B(t)$ be curves in G with $A(0) = B(0) = I$, $A'(0) = X$ and $B'(0) = Y$.

(a) Consider $C(t) = A(t)B(t)$. $C'(0) = A'(0)B(0) + A(0)B'(0) = X + Y$ so $X + Y$ is a direction. If $r \in R$, consider $D(t) = A(rt)$. $D'(0) = rA'(0) = rX$ so rX is $\hat{G}$ is a vector space. If $G \subset Gl(n)$ then $\hat{G}$ is a subspace of $M(n)$ so $\dim \hat{G} \leqq n^2$ which proves that $\hat{G}$ is finite dimensional. (Note: With more work it can be shown that $\dim \hat{G}$ is the topological dimension of G.) (b) is immediate. ■

Before presenting some examples, we recall that if $B \in M(n)$ then $\exp(B)$ is defined to the usual power series $\Sigma_{k=0}^{\infty} B^k/k!$. This series converges uniformly on any compact subset of $M(n)$ and so we may differentiate $\exp(tA)$ by differentiating the series term by term. Note also that if $\{\lambda_1, \ldots, \lambda_n\}$ are the eigenvalues of B then $e^{\lambda_1}, \ldots, e^{\lambda_n}$ are the eigenvalues of $\exp(B)$.

PROPOSITION 3.2. (a) *If $G = SO(n)$ then $\hat{G} = \{X \in M(n) \mid X + X^T = 0\}$ (skew symmetric matrices).* (b) *If $G = Sl(n)$ then $\hat{G} = \{X \in M(n) \mid$ trace $(X) = 0\}$.* (c) If *$G = A(2)$ then*

$$\hat{G} = \left\{ \begin{pmatrix} x_{11} & x_{12} & y_1 \\ x_{21} & x_{22} & y_2 \\ 0 & 0 & 0 \end{pmatrix} \,\middle|\, x_{ij}, y_i \in R \right\}.$$

Proof: (a) If $X \in \hat{G}$ then there is $A(t) \in SO(n)$ with $A(0) = I$ and $A'(0) = X$. Since $A(t)A(t)^T = I$ for all t we obtain $A'(t)A(t)^T + A(t)A'(t)^T = 0$ so that $A'(0)A(0)^T + A(0)A'(0)^T = 0$. Since $X = A'(0)$ and $A(0) = I$ this yields $X + X^T = 0$. If on the other hand we are given X which is skew-symmetric, consider $A(t) = \exp tX$. Certainly $A(0) = I$ and $A'(0) = X$ (differentiate term by term). $\det A(t) = 1$ for all t since the trace of X is zero. (See part (b) of the proof.) We need only show that $A(t)$ is an orthogonal matrix for all t.

$$A(t)A(t)^T = \left(\sum \frac{(tX)^k}{k!}\right)\left(\sum_l \frac{(tX^T)^l}{l!}\right) = \sum \frac{(tX)^k}{k!} \sum \frac{(-1)^l (tX)^l}{l!} = I,$$

where the last equality is obtained by multiplying the two series out term by term and rearranging. $A(t)$ is therefore in $SO(n)$ for all t.

(b) If $X \in \hat{G}$ then there is $A(t) \in Sl(n)$ with $A(0) = I$ and $A'(0) = X$. Let S_n be the set of all permutations of n letters, $(-1)^\sigma$ be the sign of $\sigma \in S_n$ and write $A(t) = (a_{ij}(t))$. Now $\det A(t) = \Sigma_{\sigma \in S_n} (-1)^\sigma a_{1\sigma(1)}(t) \ldots a_{n\sigma(n)}(t)$ so that

$$(\det A(t))'(0) = \sum_{k=1}^{n} \sum_{\sigma \in S_n} (-1)^\sigma a_{1\sigma(1)}(0) \ldots a_{k\sigma(k)}'(0) \ldots a_{n\sigma(n)}(0).$$

The point now is that $a_{ij}(0) = 0$ if $i \neq j$ $(A(0) = (a_{ij}(0)) = I)$ so that every term in the interior sum is

zero, except when σ is the identity permutation. We now have

$$(\det A(t))'(0)=\sum_{k=1}^{n} a_{11}(0)\dots a'_{kk}(0)\dots a_{nn}(0)=\sum_{k=1}^{n} a'_{kk}(0)=\operatorname{trace}(X).$$

Because $A(t)\in Sl(n)$, $\det A(t)=1$ so that $(\det A(t))'=0$ hence trace $(X)=0$.

Conversely if trace $(X)=0$ we show that $X\in\hat{G}$. Again consider $A(t)=\exp tX$, then $A(0)=I$, $A'(0)=X$ so we need only show that $\det A(t)=1$. If $\lambda_1,\dots,\lambda_n$ are the eigenvalues of X then the eigenvalues of $A(t)$ are $e^{t\lambda_1},\dots,e^{t\lambda_n}$ hence $\det A(t)=e^{t\lambda_1}e^{t\lambda_2}\cdots e^{t\lambda_n}=\exp(t(\lambda_1+\cdots+\lambda_n)=e^{t\operatorname{trace}(X)}=1$.

(c) If $A(t)$ is a curve in the Affine group then (remembering the identification of $A(n)$ with a subset of $Gl(n+1)$ as in section 2) we may write

$$A(t)=\begin{pmatrix} a_{11}(t) & a_{12}(t) & b_1(t)\\ a_{21}(t) & a_{22}(t) & b_2(t)\\ 0 & 0 & 1\end{pmatrix}$$

A differentiation shows that $X=A'(0)$ must have the form

$$X=\begin{pmatrix} x_{11} & x_{12} & y_1\\ x_{21} & x_{22} & y_2\\ 0 & 0 & 0\end{pmatrix}$$

as claimed. Conversely if X has the above form then so will $(tX)^k$ for all $k>0$ (i.e., the last row will always be zero), hence $\exp tX=I+\Sigma_{k=1}^{\infty}(tX)^k/k!\in A(2)$. Since $\exp tX$ is a curve in $A(2)$ we are done. ∎

It is clear from the above examples that a very important curve in G is $\exp tX$ if $X\in\hat{G}$. The precise assumption that we make on G is that $\exp tX\in G$ for all t if $X\in\hat{G}$. We have already verified this in the three cases we shall use to obtain the classical geometries. It is true if G is an open or a closed subgroup of $Gl(n)$. More abstractly, all of this material holds if G is a Lie group, in which case $\hat{G}$ is its Lie algebra. For details about what happens in this setting see **[8]**.

How do we go about defining the set of directions at $\bar{e}\in G/L$? Intuitively G/L is G with elements which differ by L identified. Thus the directions at $\bar{e}$ should be directions of G at e with elements which differ by $\hat{L}$ identified. Since $\hat{L}$ is a vector subspace of $\hat{G}$ we may form $\hat{G}/\hat{L}$ and this will be the set of directions at $\bar{e}$. Note that the set of directions at $\bar{e}$ is intrinsic. It is standard linear algebra that if $\hat{G}=\hat{L}\oplus\mathcal{P}$ for some subspace $\mathcal{P}$ of $\hat{G}$ then $\hat{G}/\hat{L}$ is isomorphic to $\mathcal{P}$. There are, however, many choices for this $\mathcal{P}$. It is the choice of a subspace $\mathcal{P}$ which will be the geometry in our situation. We will explain the second condition of the following definition and give examples later. A *geometry* on G/L is a vector subspace $\mathcal{P}$ of $\hat{G}$ such that (1) $\hat{G}=\hat{L}\oplus\mathcal{P}$ and (2) if $X\in\mathcal{P}$ and $l\in L$ then $lXl^{-1}\in\mathcal{P}$. (Here lXl^{-1} means matrix multiplication.)

What should a "line" in the geometry be? In the three examples given above if $X\in\hat{G}$ there is one very nice curve in G which X determines, $\exp(tX)$. We need only project these curves down to G/L via $\pi: G\to G/L$ to get lines. A *line* in the geometry determined by $\mathcal{P}$ is any curve in G/L defined by $A(t)=\pi(g\exp(tX))$ where $X\in\mathcal{P}$ and $g\in G$ is fixed. Note that $A(t)$ goes through $\bar{g}=gL\in G/L$. We will say that $X\in\mathcal{P}$ is tangent to $A(t)=\pi(g\exp(tX))$. Note that we have defined lines so that for each $g\in G$ $T_g: G/L\to G/L$ carries a line through xL to a line through gxL. Therefore the "group of motions" G does indeed preserve the geometric structure. The next theorem shows that each direction uniquely determines a straight line and includes in it the reason we need the second condition in the definition of geometry.

THEOREM 3.3. *Let $\mathcal{P}$ be a geometry on G/L. For each $X\in\mathcal{P}$ there is a unique line through $\bar{e}$ whose tangent is X.*

Proof: We first show that if $A(t)$ is a line through $\bar{e}$ then $A(t)$ takes the form $A(t)=\pi(\exp tX)$ for some $X\in\mathcal{P}$. If $A(t)=\pi(l\exp tX_1)$ then $l\in L$ since $A(0)=\pi(l)=\bar{e}$. Let $X=lX_1l^{-1}$ which is in $\mathcal{P}$ because of the second condition of the definition. Because $l\in L$, $\pi(l\exp tX_1)=\pi(l(\exp tX_1)l^{-1})$ so

$$A(t)=\pi(l\exp tX_1l^{-1})=\pi\left(l\sum_{k=0}^{\infty}\frac{(tX_1)^k}{k!}l^{-1}\right)$$
$$=\pi\left(\sum_{k=0}^{\infty}\frac{t^k(lX_1l^{-1})^k}{k!}\right)=\pi(\exp(tlX_1l^{-1}))=\pi(\exp X)$$

as desired.

Now suppose that $A_i(t)=\pi(l_i\exp tX_i)$ for $i=1,2$ are two lines through $\bar{e}$ and $A_1(t)=A_2(t)$. The preceding paragraph shows that we may find $\bar{X}_i\in\mathcal{P}$ such that $A_i(t)=\pi(\exp t\bar{X}_i)$. $A_1(t)=A_2(t)$ means that $C(t)=(\exp t\bar{X}_1\exp-t\bar{X}_2)\in L$ for all t. Now $C'(0)=\bar{X}_1-\bar{X}_2$ so that $\bar{X}_1-\bar{X}_2\in\hat{L}$. However, $X_i\in\mathcal{P}$ so that $\bar{X}_1-\bar{X}_2\in\mathcal{P}\cap\hat{L}=(0)$. Thus $\bar{X}_1=\bar{X}_2$ and so one line cannot represent two different directions. ■

We shall now put a geometry on each of the three homogeneous spaces S^2, H^2 and R^2 (see Table 2 and Proposition 3.2.) which will give the classical spherical, hyperbolic and Euclidean geometries and examine these in terms of the parallel postulate of Euclidean geometry.

Example 1: Recall that $G=SO(3)$,

$$L=SO(2)=\left\{\begin{pmatrix}1&0&0\\0&\multicolumn{2}{c}{\multirow{2}{*}{\boxed{B}}}\\0&&\end{pmatrix}\,\middle|\,B\in SO(2)\right\}$$

G/L is homeomorphic to S^2 by the map $\bar{\psi}(AL)=Ae$. $\hat{G}$ is the skew-symmetric matrices and

$$\hat{L}=\left\{\begin{pmatrix}0&0&0\\0&\multicolumn{2}{c}{\multirow{2}{*}{\boxed{X}}}\\0&&\end{pmatrix}\,\middle|\,X\text{ is skew-symmetric}\right\}$$

A very natural candidate for a geometry is

$$\mathcal{P}=\left\{\begin{pmatrix}0&c&-d\\-c&0&0\\d&0&0\end{pmatrix}\in\hat{G}\right\}$$

Certainly $\hat{G}=\hat{L}\oplus\mathcal{P}$ and a routine computation shows that $l\mathcal{P}l^{-1}\subset\mathcal{P}$. With this geometry S^2 is called the *Riemann Sphere*. We now find the lines of $\mathcal{P}$ through $\bar{e}$. If

$$X=\begin{pmatrix}0&1&0\\-1&0&0\\0&0&0\end{pmatrix}\text{ then }X^{2k+1}=(-1)^kX\text{ and }X^{2k}=\begin{pmatrix}(-1)^k&0&0\\0&(-1)^k&0\\0&0&0\end{pmatrix}$$

for all k. Therefore

$$\exp(tX)=\sum\frac{(tX)^k}{k!}=I+X\left(t-\frac{t^3}{3!}+\frac{t^5}{5!}-\cdots\right)+X^2\left(\frac{t^2}{2!}-\frac{t^4}{4!}+\frac{t^6}{6!}-\cdots\right)$$
$$=I+X\sin t+X^2(1-\cos t)$$

so

$$\exp tX=\begin{pmatrix}\cos t&\sin t&0\\-\sin t&\cos t&0\\0&0&1\end{pmatrix}.$$

$A(t)=\pi(\exp tX)$ is the line on G/L determined by X. The realization of this line on S^2 is

$$\bar{\psi}(A(t))=(\exp tX)(e)=\begin{pmatrix}\cos t & \sin t & 0\\ -\sin t & \cos t & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix}=\begin{pmatrix}\cos t\\ -\sin t\\ 0\end{pmatrix}.$$

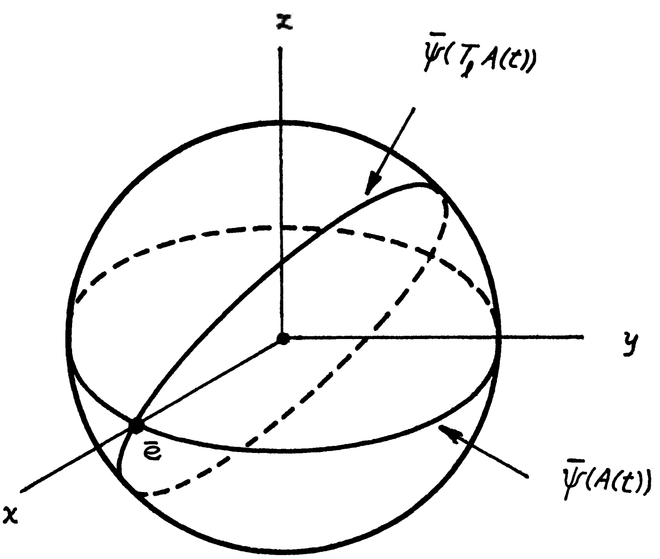


FIG. 1: Lines on the Riemann Sphere

The graph of this curve is a great circle about the equator of S^2. To find the other lines through $\bar{e}$ we need not do any more analysis because if $l\in L$ then $T_l(A(t))$ is also a line. An easy computation shows that we get all great circles through e by doing this (that is, $\bar{\psi}(T_l(A(t)))$ as l runs through L gives all the great circles of S^2 through e). See Figure 1. A similar computation shows that an arbitrary line is a great circle. We therefore may use $SO(3)/SO(2)$ with this choice of $\mathcal{P}$ as a model for spherical geometry. Note that in this geometry any two lines must meet (that is, there is no line parallel to l except l itself).

Example 2: Let H^2 be the upper half plane, $G=Sl(2)$ and $L=SO(2)$. Note that

$$\hat{G}=\left\{\begin{pmatrix}a & b\\ c & -a\end{pmatrix}\,\middle|\, a,b,c\in R^3\right\},\qquad \hat{L}=\left\{\begin{pmatrix}0 & d\\ -d & 0\end{pmatrix}\,\middle|\, d\in R\right\}$$

and recall that G/L is homeomorphic to H^2 via $\bar{\psi}(AL)=Ai$. Let

$$\mathcal{P}=\left\{\begin{pmatrix}r & s\\ s & -r\end{pmatrix}\,\middle|\, s,r\in R\right\}.$$

We call H^2 with this geometry the *Hyperbolic* (or *Poincaré*) *Upper Half Plane*. Clearly $\hat{G}=\hat{L}\oplus\mathcal{P}$. If $l\in L$ then

$$l=\begin{pmatrix}a & b\\ -b & a\end{pmatrix}\quad\text{and}\quad l^{-1}=\begin{pmatrix}a & -b\\ b & a\end{pmatrix}$$

so that for

$$X = \begin{pmatrix} r & s \\ s & -r \end{pmatrix}, \quad lXl^{-1} = \begin{pmatrix} ar^2 - 2brs - as^2 & 2ars + br^2 - bs^2 \\ 2ars + br^2 - bs^2 & -ar^2 + 2brs + as^2 \end{pmatrix}$$

which is in $\mathcal{P}$ so that $\mathcal{P}$ really is a geometry. Let

$$X = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \text{ then } \exp tX = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$$

so that $\bar{\psi}(\pi(\exp tX)) = (e^t i + 0)/(0i + e^{-t}) = e^{2t}i$, which is the positive y-axis. To find the other lines through $\bar{e}$ we once again use T_l with $l \in L$. If

$$l = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

then

$$\bar{\psi}(T_l(\exp tX)) = \frac{ae^{2t}i + b}{-be^{2t}i + a} = \frac{ab - abe^{4t} + e^{2t}i}{a^2 + b^2e^{4t}}.$$

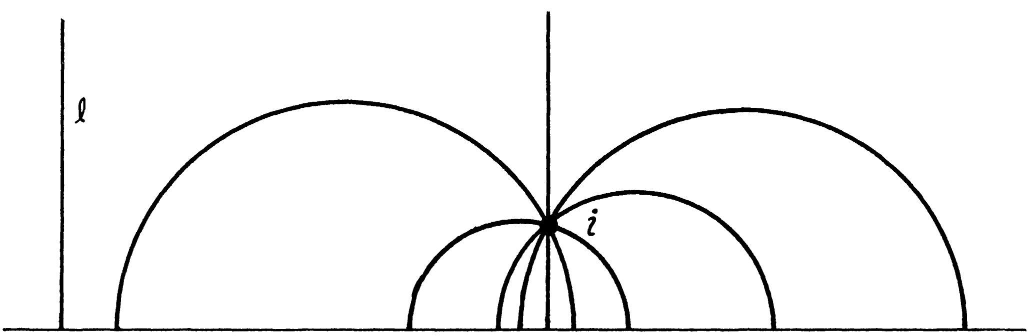


FIG. 2: Lines on the Hyperbolic Upper Half Plane

If $ab = 0$ then we get the positive y-axis again so assume $ab \neq 0$. Let $x(t) = (ab - abe^{4t})/(a^2 + b^2e^{4t})$ and $y(t) = e^{2t}/(a^2 + b^2e^{4t})$ i.e., $x(t) + iy(t) = \bar{\psi}(T_l(\exp tX))$. A monotonous computation shows that $(x(t), y(t))$ is a semicircle with center on the x-axis (at the point $((1 - 2a^2)/2ab, 0)$ and radius $(2ab)^{-1}$). Another monotonous computation shows that all the lines through $\bar{e}$ get sent by T_g (for $g \in Sl(2)$) to either a Euclidean line parallel to the y-axis or a semicircle centered on the x-axis (see Figure 2). Note that in this geometry, given a line (say l in Figure 2) and a point (say i) not on the line, there are infinitely many lines which do not meet (are "parallel to") the given line. $Sl(2)/SO(2)$ with this choice of $\mathcal{P}$ serves as a model for hyperbolic geometry.

Example 3: Recall that if $G = A(2)$, $L = Gl(2)$, G/L is homeomorphic to R^2 by the map $\bar{\psi}(AL) = A\mathbf{0}$. By Proposition 3.2c,

$$\hat{G} = \left\{ \begin{pmatrix} x_{11} & x_{12} & y_1 \\ x_{21} & x_{22} & y_2 \\ 0 & 0 & 0 \end{pmatrix} \middle| \; x_{ij}, y_i \in R \right\} \text{ and } \hat{L} = \left\{ \begin{pmatrix} x_{11} & x_{12} & 0 \\ x_{21} & x_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix} \middle| \; x_{ij} \in R \right\}.$$

A natural candidate for a geometry $\mathcal{P}$ is

$$\mathcal{P} = \left\{ \begin{pmatrix} 0 & 0 & a \\ 0 & 0 & b \\ 0 & 0 & 0 \end{pmatrix} \middle| \; a, b \in R \right\}.$$

It is easy to check that $\mathscr{P}$ is a geometry. If $X \in \mathscr{P}$ then $X^k = 0$ for all $k \geqq 2$ so that

$$\exp tX = I + tX = \begin{pmatrix} 1 & 0 & ta \\ 0 & 1 & tb \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \boxed{|I|} & & ta \\ & & tb \\ 0 & 0 & 1 \end{pmatrix}.$$

If $A(t) = \pi(\exp tX)$ then

$$\bar{\psi}(A(t)) = \begin{pmatrix} \boxed{I} & & ta \\ & & tb \\ 0 & 0 & 1 \end{pmatrix} 0 = (ta, tb).$$

Thus $A(t)$ is a Euclidean straight line through the origin. It is easy to see that $T_g(A(t))$ is also a Euclidean straight line for all $g \in G$. $A(2)/Gl(2)$ with this geometry is called *Euclidean geometry.*

TABLE 3: Models for Classical Geometries

Example	Coset Space	Number of Parallels
S^2	$SO(3)/SO(2)$	None
H^2	$Sl(2)/SO(2)$	Infinitely many
R^2	$A(2)/Gl(2)$	One

This paper is an exposition of an invited address to the Fifty-fourth Annual Meeting of the Illinois Section of the MAA which was held at Rockford College, May 9 and 10, 1975. The paper was given under the title "A Non-synthetic Approach to Transformation Geometry."

References

1. L. Ahlfors, Complex Analysis, McGraw-Hill, New York, 1953.
2. B. L. Beers and A. Dragt, New theorems about spherical harmonic expansions and $SU(2)$, J. Math. and Phys., 11 (1970) 2313–2328.
3. C. Boyer, A History of Mathematics, Wiley, New York, 1968.
4. J. Coolidge, A History of Geometrical Methods, Clarendon, Oxford, 1940.
5. D. Gans, Transformations and Geometries, Appleton-Century-Crofts, New York, 1969.
6. S. Helgason, Differential Geometry and Symmetric Spaces, Academic Press, New York, 1962.
7. F. Klein, A Comparative Review of Recent Researches in Geometry, Bulletin of the New York Mathematical Society (now Bulletin of the American Mathematical Society), (1893) 215–249.
8. S. Kobayashi and K. Nomizu, Foundations of Differential Geometry, Vol. I and II. Interscience, New York, 1963 and 1969.
9. Y. Matsushima, Differentiable Manifolds, Marcel Dekker, New York, 1972.
10. R. Millman and A. Stehney, The Geometry of Connections, this MONTHLY, 80 (1973) 475–500.
11. M. Spivak, Differential Geometry, Vol. II, Chapter IVA Published by Michael Spivak, 1970.
12. A. Tuller, A Modern Introduction to Geometries, Van Nostrand, Princeton, N.J., 1967.

DEPARTMENT OF MATHEMATICS, SOUTHERN ILLINOIS UNIVERSITY, CARBONDALE, IL 62901

GROUP DECISION DEVICES

LESTER E. DUBINS

1. Introduction. This article is concerned with the ever recurring and fundamental problem of how a group of individuals can, by a symmetric or democratic procedure with due attention to individual preferences, come to make a group decision with which the individual members are reasonably satisfied. Fundamental results of a negative nature have been given by Arrow **[1]**, and others, and a few of a positive result, especially for a group consisting of two people, by the theory of two-person, zero-sum games **[5]** and **[6]**. In another direction, an important positive contribution is discussed by Steinhaus in **[11]** who attributes the useful ideas presented there to S. Banach and M. Knaster. This paper offers a device, similar in spirit to that of Knaster, but wider in scope. Under the device suggested below, the group decision arrived at is accompanied by a system of transfers of cash among the members of the group, which transfers, roughly speaking, can be viewed as a means by which those who are relatively content with the decision compensate those who are relatively discontent. The device is offered as a practical means for settling some group decision problems. Illustrative examples are provided in Sections 8 and 9 which clarify the general theory and which some readers may wish to refer to early in their reading of this article.

2. Decision problems, functions, and procedures. A *decision problem* is an ordered pair (G, D) of nonempty, finite sets G and D. Informally, G is the group of people who must choose one, and only one, decision from the set D of possible decisions.

A *choice function* is any rule C which assigns to each $g \in G$ an arbitrary nonempty set $C(g)$. Each element of $C(g)$ is to be thought of as a possible choice for g. Let the set ΠC be the Cartesian product of the sets $C(g)$ for $g \in G$ and call ΠC the set of possible *plays* or *elections*. That is, an element of ΠC is simply a function x defined on G for which $x(g) \in C(g)$ for all $g \in G$. Such an x is called a *play* or an *election*, and $x(g)$ is g's *vote*, *choice*, or *pure strategy* in the election x.

A *decision function* (for C, D) is a D-valued function, δ, defined on ΠC. Call (C, δ) a *decision procedure* for the decision problem (G, D). For some decision problems (G, D), it is reasonable to hope that the group, though unable to agree on a decision directly, can agree to use a specific decision procedure (C, δ) to arrive at a decision, thus. Each $g \in G$ chooses some $x(g) \in C(g)$. This system of choices, that is, this election, x, together with the rule δ, then determines the decision, $\delta(x)$.

As here defined, decision functions are very wide in scope. They include the social welfare functions as defined in Arrow **[1**, pp. 22], which becomes evident by letting $C(g)$ be, for each $g \in G$, the set of all linear orderings of D.

As certain work of Arrow **[1]** suggests, even decision procedures offer only limited hope of achieving decisions even moderately satisfactory to all. In order to achieve more often decisions that are satisfactory to all members of the group, the decision function will be augmented by a system of cash transfers, as will now be explained.

3. Decision devices. A *cash exchange* function is a real-valued function c, defined on $G \times \Pi C$, which satisfies

$$\sum_{g \in G} c(g, x) = 0 \quad \text{for all } x \in \Pi C. \tag{1}$$

The interpretation of $c(g, x)$ is that it is the cash that is given to g in the event of the election x, where the unit of cash is, say, a dollar. (If $c(g, x)$ is negative, then $|c(g, x)|$ is the cash that g contributes to the other members of G.)

A *decision device* is an ordered couple (δ, c) where δ is a decision function and c is a cash exchange function, both functions based on the same choice function, C.

What does it mean for the members of G to agree to the use of the device (δ, c) to resolve a decision problem (G, D)? It means that each $g \in G$ agrees to make a choice $x(g)$ from $C(g)$ (which choice is perhaps best revealed to the other members of G only after they, too, have made their choices). The resulting election, x, together with the decision rule, δ, determines the group decision, $\delta(x)$ as well as the amount of cash, $|c(g, x)|$, that g receives or disburses — receives if $c(g, x)$ is positive, disburses if $c(g, x)$ is negative.

4. Attractive decision devices. Let R be the set of real numbers, R^D the finite-dimensional vector space of all R-valued functions defined on D, and H the subspace of R^D consisting of all $y \in R^D$ for which

$$\sum_{d \in D} y(d) = 0. \tag{1}$$

Henceforth, let $C(g) = H$ for all $g \in G$. So each play $x \in \Pi C$ determines, and can be identified with, a rectangular matrix of real numbers

$$x(g, d) = x(g)(d), \quad g \in G, d \in D, \tag{2}$$

with each row vector an element of H, that is, with

$$\sum_{d \in D} x(g, d) = 0 \quad \text{for each } g \in G. \tag{3}$$

A decision device (δ, c) is *attractive to* g provided that, for each $y \in H$, and each play x, if $x(g) = y$, then

$$c(g, x) \geqq y(\delta(x)). \tag{4}$$

In words, the device (δ, c) is attractive to g if, upon its employment by the group, he can set a lower bound, $y(d)$, to the cash that he is to receive in the event that the decision is d, where the lower bound $y(d)$ is arbitrary, subject only to the constraint (1). If the device (δ, c) is attractive to every $g \in G$, the device is *attractive*. For decision problems for which an attractive device has been agreed upon, it is suggestive to call $-x(g)$, g's *offer*. There plainly exist attractive devices and, in fact, the set of them can easily be described. Notice that for each play, x, there is a $d \in D$ for which

$$\sum_{g \in G} x(g, d) \leqq 0, \tag{5}$$

as is evident in view of (3). Call δ *admissible* if, for every play x,

$$\sum_{g \in G} x(g, \delta(x)) \leqq 0. \tag{6}$$

Plainly, there exist admissible δ. So let δ be admissible and let e be a nonnegative function such that

$$\sum_{g \in G} e(g, x) = -\sum_{g \in G} x(g, \delta(x)) \tag{7}$$

for all x. Finally, let

$$c(g, x) = x(g, \delta(x)) + e(g, x). \tag{8}$$

As is easily verified, (δ, c) is then an attractive decision device, and every attractive decision device is of this form for some admissible δ.

5. The consideration of utility. Here a decision device (δ, c) is seen from the point of view of some particular member g of G. Let $u(d, c)$ be the utility that g ascribes to the joint event that the

group decision is d and that he receives a cash payment of c dollars. Hence, for an election x which leads to the decision $\delta(x)$ and a cash transfer to g of $c(g,x)$, its utility to g, say, $u(x)$, is given by

$$u(x) = u(\delta(x), c(g,x)) \\ = u(\delta(x), c(x)), \tag{1}$$

where the dependence on g is temporarily not indicated.

How high can g guarantee $u(x)$ to be no matter what the choices of the other members of G? Let $y \in C(g) = H$, and let

$$J(y) = \inf u(x), \tag{2}$$

where the infimum is taken over all x for which $x(g) = y$. Let

$$v = \sup_{y \in H} J(y). \tag{3}$$

Then by appropriate choice of his pure strategy, g can be certain of an outcome whose utility to him is at least arbitrarily close to v, but no higher utility can be guaranteed by g alone, which justifies labelling v as g's *lower value.*

The determination of v made here will involve the following relatively innocuous assumption, which is henceforth in force: *For each $d \in D$, $u(d,c)$ is continuous and strictly increasing in c.*

LEMMA 1. *For all attractive decision devices,*

$$J(y) = \inf_{d \in D} u(d, y(d)). \tag{4}$$

Proof: Let z be the sûm of the pure strategies of all members of G other than g, and consider a $z \in H$ which equals $-y + \varepsilon_d$, where $\varepsilon_d(d) = -\varepsilon$ and $\varepsilon_d(d') \geqq 0$ for all other d'. Then the election x has only one column whose sum is negative, namely that corresponding to d, so $\delta(x) = d$ and

$$y(d) \leqq c(g,x) \leqq y(d) + \varepsilon, \tag{5}$$

as is easily verified. Hence,

$$u(x) = u(\delta(x), c(g,x)) \\ = u(d, c(g,x)) \\ \leqq u(d, y(d) + \varepsilon), \tag{6}$$

where the inequality holds in view of the assumed monotoneity of u in c. Consequently,

$$J(y) \leqq u(d, y(d) + \varepsilon), \tag{7}$$

and, since (7) holds for all $\varepsilon > 0$, and $u(d,c)$ is assumed continuous in c, (7) holds also for $\varepsilon = 0$. Since d is arbitrary, $J(y)$ is, therefore, at most the right-hand side of (4). For the reverse inequality, calculate thus:

$$\varepsilon + J(y) \geqq u(x) = u(\delta(x), c(g,x)) \\ \geqq u(\delta(x), y(\delta(x))) \\ \geqq \inf u(d, y(d)). \tag{8}$$

This completes the proof of Lemma 1.

LEMMA 2. *Let $u(d,c)$ be a strictly increasing and continuous function of a real variable c, for d ranging*

over a finite set D. Then there is one, and only one, (y^*, u^*) *such that* y^* *is a real-valued function defined on D for which* $\Sigma_d y^*(d) = 0$, *and* u^* *is a real number for which*

$$u(d, y^*(d)) = u^* \quad \text{for all } d \in D. \tag{9}$$

If moreover, $u(d, 0)$ *is in the range of* $u(d', 0)$ *for each d and d', in D, then*

$$u^* = \sup_{y \in H} \inf_{d \in D} u(d, y(d)). \tag{10}$$

Proof. For each d, $u(d, \cdot)$ has an inverse function $u^{-1}(d, \cdot)$ where

$$u(d, c) = t \leftrightarrow u^{-1}(d, t) = c. \tag{11}$$

So

$$u(d, y(d)) = t \leftrightarrow u^{-1}(d, t) = y(d). \tag{12}$$

Hence the conclusion of the first part of the lemma is equivalent to the assertion that there exists a unique real number t such that

$$\sum_{d \in D} u^{-1}(d, t) = 0. \tag{13}$$

That there is at most one such t is clear since each u^{-1} is strictly increasing in t. In view of the continuity of $u^{-1}(d, t)$ in t, the existence of a t which satisfies (13) will be seen once it is seen that the left-hand side of (13) assumes both positive and negative values. Plainly, $u^{-1}(d, t)$ is negative for t less than $u(d, 0)$ and positive for t greater than $u(d, 0)$. Consequently, for every d, $u^{-1}(d, t)$ is negative to the left of $\min_d u(d, 0)$ and positive to the right of $\max_d u(d, 0)$. Hence the left-hand side of (13) does assume both negative and positive values, which completes the proof of existence and uniqueness. To see that u^*, namely the solution t above, is no greater than the right-hand side of (10), note that for $y = y^*$, $\inf_d u(d, y(d))$ is u^*. For the reverse inequality, let y be any element of H, verify that, for at least one d, $y^*(d) \geqq y(d)$, and compute thus:

$$\begin{aligned} u^* &= u(d, y^*(d)) \\ &\geqq u(d, y(d)) \\ &\geqq \inf_{d \in D} u(d, y(d)), \end{aligned} \tag{14}$$

which completes the proof.

In view of Lemmas 1 and 2, the following is evident.

PROPOSITION 1. *For all attractive decision devices and all* $u(d, c)$, *as in Lemma* 2, u^* *is the lower value,* v, *and the sup in* (3) *is attained at, and only at,* $y = y^*$.

The same technique for determining v which is a $\sup\inf u(x)$ applies to determine $\inf\sup u(x)$, and the value is again u^*.

It is noteworthy that in contrast to the sup in (3) being attained, the inf in (2) need not be attained. This lack of attainment is due to discontinuities in $u(x)$, as a function of x, which in turn is due to inescapable discontinuities in $\delta(x)$.

6. The case of utilities which are linear in cash. Continue to fix a particular $g \in G$. Call his utility u *linear* if it has this special form

$$u(d, c) = u(d) + c, \tag{1}$$

where, of course, $u(d)$ is then interpretable as the cash utility to him of the decision, d.

Call the y^* of Lemma 2 and Proposition 1 the *max-min strategy* of g.

PROPOSITION 2. *If the decision device is attractive and g's utility is linear, as in* (1), *then g's lower value v is given by*

$$v = \frac{1}{m} \sum_{d \in D} u(d), \tag{2}$$

where m is the cardinality of D, and g's max-min strategy y^* is given by

$$y^*(d) = -u(d) + \frac{1}{m} \sum_{d \in D} u(d). \tag{3}$$

Proof: According to (9) of the preceding section,

$$\begin{aligned} u^* &= u(d, y^*(d)) \\ &= u(d) + y^*(d) \quad \text{for all } d \in D. \end{aligned} \tag{4}$$

Hence,

$$mu^* = \sum_{d \in D} u(d) + \sum_{d \in D} y^*(d) = \sum_{d \in D} u(d) + 0, \tag{5}$$

where m is the cardinality of D, so u^* equals the right-hand side of (2). In view of Proposition 1, so does v. Now (3) plainly follows from (4).

7. Meritorious decision devices. Let $e(x, d)$ be $-\Sigma x(g, d)(g \in G)$, and, somewhat suggestively, call $e(x, d)$ the *bonus* if the election x and the decision d occur. Let

$$e(x) = \sup_{d \in D} e(x, d) = -\inf_{d \in D} \sum_{g \in G} x(g, d), \tag{1}$$

and say that δ is *meritorious* if, for every x, $\delta(x)$ is some d for which the inf in (1) is attained, that is, some d for which

$$e(x) = -\sum_{g \in G} x(g, d) = -\sum_{g \in G} x(g, \delta(x)). \tag{2}$$

Certain cash transfer functions c are mathematically tractable. Say that the decision device (δ, c) is *meritorious* if δ is meritorious and if

$$c(g, x) = x(g, \delta(x)) + \frac{1}{n} e(x), \tag{3}$$

where n is the cardinality of G. Plainly, meritorious devices are necessarily attractive. Henceforth, none but meritorious devices will be treated.

Call the election x^* that results when every member of G uses his max-min strategy the *max-min election*, and suppose henceforth that every $g \in G$ has a linear utility $u(g, d, c)$,

$$u(g, d, c) = u(g, d) + c. \tag{4}$$

For such utilities there may be some justification in calling

$$U(d) = \sum_{g \in G} u(g, d) \tag{5}$$

the *social utility* of decision d. Thus, under any election x, the decision $\delta(x)$ has a certain social utility $U(\delta(x))$. If (δ, c) is meritorious and every g uses his max-min strategy, then the resulting election has the maximum possible social utility, as is easily verified.

There is extant a considerable literature dealing with game theory, utility theory, and other topics that the present article touches upon. For the reader interested in pursuing known results about these matters, classical as well as recent, a list of references is appended.

8. On sharing the indivisible — An inheritance illustration. A man's will declares that his three sons, named 1, 2, and 3, are to share equally in his estate which turns out to consist of a single painting. Let $u(j)$ be the utility to, say, the first son, of the decision d_j that the painting be given to j. Suppose, for example, that $u(j)$ is 300 or 0 according as j is 1 or not. Then 1's max-min strategy y^* is (200, -100, -100), which, if he uses it, assures him of an effective inheritance of at least 100, namely, either 100 in cash, or the painting, which to him is worth 300, at a cost of at most 200. This outcome is identical with that obtained by the method attributed by Steinhaus in **[11]** to Knaster.

If, however, son 1 would prefer that the painting be given to one of his brothers rather than to the other, the method proposed by Knaster does not offer the possibility for him to express such a preference, whereas, even for linear utilities, meritorious decision devices do. For instance, if the first son's choice of strategy y is (200, -50, -150), then some one of these three possibilities must occur: (i) he receives the painting at a cash cost to him of at most 200; (ii) the painting goes to the second brother and 1 receives at least 50; or, (iii) the painting goes to the third brother and 1 receives at least 150.

Another contrast between Knaster's method and devices of this note will now be illustrated. Suppose that, in addition to the painting, there is a piece of sculpture in the estate. The Knaster method, being equivalent to that of placing separate bids for the sculpture and painting, does not provide an opportunity for an heir to indicate whether his possession of both art items is worth more (or less) to him than possession of each separately. In contrast, those decision devices called attractive in this note do. To see this, let D consist of 9 decisions corresponding to the nine different ways in which the two objects can be assigned to the three brothers, so a strategy is any vector y with nine components subject to the constraint that their sum is zero.

9. An illustration concerning the election of a reluctant department chairman. Suppose that a department needs a chairman for next year and that the chairmanship is considered somewhat onerous by all 51 faculty members of the department, one of whom must serve. Suppose, too, that all feel a responsibility to share in this duty and that they therefore agree to participate in a meritorious decision device described above. If, for example, a member feels that his serving as chairman has a negative utility to him of about \$5100, he could choose to bid 100 for each of his 50 colleagues, and -5000 for himself. In this way, either he contributes at most 100 in cash to the colleague who serves, or else he serves, and receives in compensation at least 5000, so that again his net contribution is, in effect, at most 100. If the chairman is selected by the dean, by lottery, or by an ordinary election, some one individual bears a burden equivalent to about a \$5100 tax for the year, and all other members contribute nothing. In contrast, under the meritorious decision device, the burden is spread rather uniformly over the entire department. Moreover, suppose that the chairmanship is particularly odious to an individual, g, and/or suppose g especially prefers some particular member or members of the department to be or not to be chairman. Then under the decision devices introduced above, g's strategy can reflect this, which not only correspondingly increases or decreases the prospect of this or that member from becoming chairman, but also assures g of appropriately varying minimal cash compensation (or of maximal cash cost). If engaging in exchanges of cash has too great a negative moral or emotional connotation, bids could be made in terms of some other commodity, for instance hours of additional, or diminished, class instruction time.

The two illustrations suggest that the meritorious decision device might play a role in the resolution of real economic and political problems, possibly including international disputes.

Research supported by National Science Foundation Grant GP–43085.

References

1. Kenneth J. Arrow, Social Choice and Individual Values, Wiley, New York, 1951; 2nd ed., Cowles Foundation Monograph 12, Yale University Press, New Haven, 1963.
2. L. E. Dubins and E. H. Spanier, How to cut a cake fairly, this MONTHLY, 68 (1961, Part I) 1–17.
3. Peter C. Fishburn, Utility Theory for Decision Making, Wiley, New York, 1970.
4. Harold W. Kuhn, On games of fair division, Essays in Mathematical Economics in Honor of Oskar Morgenstern, edited by Martin Shubik, Princeton University Press, Princeton, N.J., 1967.
5. R. D. Luce and H. Raiffa, Games and Decisions, Wiley, New York, 1957.
6. J. von Neumann and O. Morgenstern, Theory of Games and Economic Behavior, Princeton University Press, Princeton, 3rd ed., 1953.
7. Leonard J. Savage, The Foundations of Statistics, Wiley, New York, 1954; 2nd ed., Dover, New York, 1972.
8. Thomas C. Schelling, The Strategy of Conflict, Harvard University Press, Cambridge, Mass., 1960.
9. Amarta K. Sen, Collective Choice and Social Welfare, Holden-Day, San Francisco, 1970.
10. Lloyd S. Shapley and Martin Shubik, (forthcoming book) Game Theory in Economics, Chapter 4: "Preferences and Utility"; presently, Report R–904/4–NSF, RAND Corporation, Santa Monica, 1974.
11. H. Steinhaus, Sur la division pragmatique, Econometrica (supplement), 17 (1949) 315–319.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CA 94720.

SOLUTION OF WARING'S PROBLEM MOD n

CHARLES SMALL

This paper is an addendum to [1] in which we introduced the numbers $g(k,n)$, defined for all integers k and n greater than 1 as the smallest r such that every integer is a sum of r kth powers mod n.

By "*Waring's problem mod* n" we mean the problem of computing $g(k,n)$ as a function of k and n; more precisely, we seek for each fixed k, a theorem which computes $g(k,n)$ for all n, or which at least reduces the computation for *all* n to the computation for an explicitly identified *finite* set of n.

For $k=2$ and $k=3$ we gave complete solutions in [1], Sections 3 and 5 respectively. In this sequel, we first give complete solutions for $k=4$ and $k=5$, and then observe that the same techniques yield a solution for *all* k, in the form of a theorem reducing the computation for *all* n to the computation for an *explicit finite* list of n (depending on k). The cases when k is even or odd are treated separately, but the resulting theorems are similar.

To begin, fix k and consider $g(k,p)$ for prime p. Let $l=(k,p-1)$; then $g(k,p)=g(l,p)$, and $g(l,p)=l$ if $l=p-1$ or $(p-1)/2$ ([1], 4.2). We recall also from [1] that when p is prime and $p>(k-1)^4$ we have $g(k,p)\leqq 2$, with equality if, and only if, $(k,p-1)>1$. (The " $=$ " in the assertion "$g(k,p)=2$ if $p>(k-1)^4$", at the top of the list of inequalities near the end of Section 4 of [1], should be " $\leqq$ ". The criterion for equality comes from [1], 4.1: $(k,p-1)=(G\colon G^k)$, where G is the group of units mod p.) Thus, we know $g(k,p)$ for all primes p if we know $g(l,p)$ for the proper divisors l of k and $g(k,p)$ for those $p\equiv 1 \pmod k$ satisfying $k\leqq(p-1)/3$ and $p<(k-1)^4$, or in other words, $3k+1\leqq p<(k-1)^4$. For each such p, $g(k,p)$ is one of the numbers $2,3,\ldots,[k/2]+1$ ([1], 4.3). For $k=4$ and $k=5$ there are eight and twelve primes (respectively) to worry about; the results for these values are shown in the table below. Thanks are due to L. G. Roberts who programmed the computations.

p	13	17	29	37	41	53	61	73
$g(4,p)$	3	3	3	2	3	2	2	2

p	31	41	61	71	101	131	151	181	191	211	241	251
$g(5,p)$	3	3	3	3	3	2	2	2	2	2	2	2

The next step is to lift information mod p to information mod p^e, $e>1$. This works beautifully when k is odd:

LEMMA 1. *Let k be odd, let p be prime, and define $c \geqq 0$ by: $p^c \| k$ (i.e., $p^c \mid k$, $p^{c+1} \nmid k$). Then:*
Case 1. $c=0$, $(k,p-1)>1$. *Then* $g(k,p^e)=g(k,p)$ *for all* $e \geqq 1$.
Case 2. $c=0$, $(k,p-1)=1$. *Then* $g(k,p)=1$ *and* $g(k,p^e)=2$ *for all* $e>1$.
Case 3. $c>0$. *Then* $g(k,p) \leqq g(k,p^2) \leqq \cdots \leqq g(k,p^{2c+1})=g(k,p^e)$ *for all* $e>2c+1$.

Proof. Since k is odd we can take $t(k,p)=2$ in 2.2 of [**1**], and when $(k,p-1)>1$ we have $g(k,p) \geqq 2$ ([**1**], 4.1). Thus in Case 1 the inequalities $g(k,p) \leqq g(k,p^e) \leqq \max(g(k,p), t(k,p))$ of [**1**], 2.2 yield $g(k,p)=g(k,p^e)$. In Case 2 we have $g(k,p)=1$ by [**1**], 4.1 and $g(k,p^e) \leqq 2$ for all $e>1$ by [**1**], 2.2; but $g(k,p^e) \geqq 2$ for all $e>1$ by [**1**], 5.2. The conclusion in Case 3 just restates [**1**], 2.4.

This lifting lemma gives the theorem for $k=5$:

THEOREM 1. *Let $r \geqq 1$ and $n>1$ be integers. Then every element of $\mathbb{Z}/(n)$ is a sum of r fifth powers if and only if one of the following holds:*
$r=1$ *and n is a product of distinct primes, none of them* $\equiv 1 \pmod{10}$;
$r=2$ *and n is not divisible by any of*: 11, 25, 31, 41, 61, 71, 101;
$r=3$ *and* $11 \nmid n$;
$r \geqq 5$ (*and no condition on* n).
(In particular, if four 5th powers suffice then in fact three suffice.)

To prove this, observe first that if $n=p_1^{e_1}\cdots p_s^{e_s}$ (p_i distinct primes, $e_i \geqq 1$) then $g(k,n)=\max_{1\leqq i\leqq s} g(k,p_i^{e_i})$ (Chinese Remainder Theorem). Hence Theorem 1 is equivalent to:

THEOREM 1$'$. *For any prime p and any $e \geqq 1$,*

$$g(5,p^e)=\begin{cases} 5 & \text{if } p=11 \\ 3 & \text{if } p=31,41,61,71,101 \\ 2 & \text{if } p\equiv 1 \pmod{10},\ p\geqq 131 \\ 1 & \text{if } p\not\equiv 1 \pmod{10},\ e=1 \\ 2 & \text{if } p\not\equiv 1 \pmod{10},\ p\neq 5,\ e>1 \\ 3 & \text{if } p=5,\ e>1. \end{cases}$$

Proof of Theorem 1$'$. Assume first that $p\equiv 1 \pmod{10}$. Then for $e=1$ the results follow from [**1**], 4.2 ($p=11$), the table above ($31\leqq p\leqq 251$), and the "diagonal equations" argument indicated toward the end of Section 4 of [**1**]: when $p>(k-1)^4$, $g(k,p)=2$. For $e>1$, and for $p\not\equiv 1 \pmod{10}$, $p\neq 5$, use Cases 1 and 2 of Lemma 1. Finally, for $p=5$, compute $g(5,25)=3$ directly (the fifth powers mod 25 are 0, ± 1, ± 7) and invoke [**1**], 7.9 and Case 3 of Lemma 1. (The reference to [**1**], 7.9 can of course be circumvented at the cost of computing $g(5,125)=3$ directly.)

Clearly, the same arguments compute $g(k,n)$ completely (i.e., reduce the computation for *all* n to a *finite* task) whenever k is odd:

THEOREM 2. *Let $k=p_1^{e_1}\cdots p_s^{e_s}$ (p_i distinct odd primes, $e_i\geqq 1$). Denote by $q_1,\ldots,q_t$ the primes $q<(k-1)^4$ such that $q\equiv 1 \pmod{2p_i}$ for some $i=1,\ldots,s$ and for each $l=1,\ldots,t$ put $b_l=g(k,q_l)$. For each $i=1,\ldots,s$ and $j=1,\ldots,2e_i+1$ put $a_{ij}=g(k,p_i^j)$. Then the $s+t+2\Sigma e_i$ numbers $\{b_l,a_{ij}\}$ determine $g(k,n)$ for all n. Explicitly, for any prime p and $e\geqq 1$ we have:*
Case 1. *$p\neq p_i$ for all i and $p\equiv 1 \pmod{p_i}$ for some i, $1\leqq i\leqq s$. Then if $p<(k-1)^4$, $p=q_l$ for some l ($1\leqq l\leqq t$) and $g(k,p^e)=b_l$; and if $p>(k-1)^4$, $g(k,p^e)=2$.*
Case 2. *$p\neq p_i$ and $p\not\equiv 1 \pmod{p_i}$, for all i, $1\leqq i\leqq s$. Then $g(k,p^e)=1$ if $e=1$ and 2 if $e>1$.*
Case 3. *$p=p_i$. Then $g(k,p^e)=a_{ij}$ where $j=\min(e,2e_i+1)$.*

Proof: For $\alpha = 1, 2, 3$, Case α in Theorem 2 comes from Case α in Lemma 1.

For example, $g(45, n)$ is determined for *all* n if we know it for $n = 3, 3^2, 3^3, 3^4, 3^5, 5, 5^2, 5^3$, and all primes $n < (44)^4$ such that $n \equiv 1 \pmod 6$ or $n \equiv 1 \pmod{10}$.

For even k we again consider first $g(k, p)$ for primes p. Again, the important primes p are those satisfying $p \equiv 1 \pmod k$ and $3k + 1 \leqq p < (k-1)^4$. Define $t(k, p)$ as in [1] to be the smallest s such that 0 is nontrivially a sum of s kth powers mod p. Then $t(k; p) \leqq g(k, p) + 1$ in any case, and $t(k, p) \leqq g(k, p)$ if and only if -1 is a sum of fewer than $g(k, p)$ kth powers mod p. For primes p satisfying $t(k, p) \leqq g(k, p)$ we have, (if $p \nmid k$), $g(k, p^e) = g(k, p)$ for all $e \geqq 1$, as in Case 1 of Lemma 1, from [1], 2.2. When $t(k, p) = g(k, p) + 1$ the lifting lemma is less decisive; [1], 2.2 gives only $g(k, p) \leqq g(k, p^e) \leqq g(k, p) + 1$ for all $e \geqq 1$ (again assuming $p \nmid k$). Part 3 of the following lemma rescues the situation in this case:

LEMMA 2. *Let $k = 2^a l$, $a \geqq 0$, l odd, and let p be an odd prime. Then*

(1) -1 *is a k-th power* $\bmod p$ *if, and only if,* $p \equiv 1 \pmod{2^{\alpha+1}}$.

(2) *If* $p \not\equiv 1 \pmod{2^{a+1}}$, *then* $g(k, p^e) \geqq 3$ *for all* $e > 1$.

(3) *If* $g(k, p) < t(k, p)$, *then* $g(k, p) < g(k, p^2)$; *if, in addition,* $p \nmid k$, *then* $g(k, p^e) = g(k, p) + 1$ *for all* $e > 1$.

(4) *If* $g(k, p) = 1$ *(equivalently,* $(k, p-1) = 1$*), then* $t(k, p) = 2$; *if* $g(k, p) = 2$ *(which happens, for example, whenever* $p > (k-1)^4$ *and k is even) then* $t(k, p) = 2$ *if* $p \equiv 1 \pmod{2^{a+1}}$ *and* $t(k, p) = 3$ *otherwise.*

Proof. (1) is a special case of Theorem 1 of [2]. (2) and (3) will be proved simultaneously. Suppose p is a sum of n kth powers mod p^2: $p \equiv x_1^k + \cdots + x_n^k \pmod{p^2}$. We cannot have $p \mid x_i$ for all i, for then $p \equiv 0 \pmod{p^2}$, which is absurd. Hence 0 is nontrivially a sum of n kth powers mod p, i.e., $t(k, p) \leqq n$. Taking $n = 2$ and using part (1) this argument shows that $g(k, p^2) > 2$, which gives part (2). Taking $n = g(k, p)$, it gives (the contrapositive of) the first half of part (3). The rest of (3) follows from 2.2 of [1], according to which $g(k, p^e) \leqq g(k, p) + 1$ whenever $p \nmid k$. (4) follows from (1).

To illustrate how Lemma 2 solves the problem when k is even, we give the explicit solution for $k = 4$:

THEOREM 3. *Let $r \geqq 1$ and $n > 1$ be integers. Then every element of $\mathbb{Z}/(n)$ is a sum of r fourth powers if and only if one of the following holds*:

$r = 1$ *and* $n = 2$.

$r = 2$ *and n is not divisible by any of*: 4, 5, 13, 17, 29, 41, *and* p^2 *where p is prime and* $p \equiv 3 \pmod 4$ *or* $29 < p \equiv 5 \pmod 8$.

$r = 3$ *and n is not divisible by* 5, 8, *or* $(29)^2$.

$r = 4$ *and n is not divisible by* 8 *or* 25.

$r = 5$ *and* $8 \nmid n$.

$r = 7$ *and* $16 \nmid n$.

$r \geqq 15$ *(and no condition on n).*

(In particular if fewer than fifteen 4th powers suffice, then in fact seven suffice, and if six suffice then five suffice.)

As in the other cases, we use the Chinese Remainder Theorem to translate this to the computation of $g(4, p^e)$ for primes p and $e \geqq 1$ and prove it in that form:

THEOREM 3′. *Let p be prime and* $e \geqq 1$, *then*:

$g(4, p^e) = 1$ *if* $p = 2$ *and* $e = 1$.

$g(4, p^e) = 2$ *if* $41 < p \equiv 1 \pmod 8$; *if* $p \equiv 3 \pmod 4$ *and* $e = 1$; *and if* $29 < p \equiv 5 \pmod 8$ *and* $e = 1$.

$g(4, p^e) = 3$ *if* $p = 2$ *and* $e = 2$; *if* $p = 13, 17,$ *or* 41; *if* $p \equiv 3 \pmod 4$ *and* $e > 1$; *if* $p = 29$ *and* $e = 1$; *and if* $29 < p \equiv 5 \pmod 8$ *and* $e > 1$.

$g(4,p^e)=4$ *if* $p=5$ *and* $e=1$; *and if* $p=29$ *and* $e>1$.
$g(4,p^e)=5$ *if* $p=5$ *and* $e>1$.
$g(4,p^e)=7$ *if* $p=2$ *and* $e=3$.
$g(4,p^e)=15$ *if* $p=2$ *and* $e\geqq 4$.

Proof. For $p=2$, compute $g(4,2^e)$ directly for $e\leqq 5$ (the fourth powers mod 32 are $0,1,16,17$) and use [1], 2.4 to get the result for $e>5$. Whenever $p\equiv 1 \pmod 8$ part (1) of Lemma 2 gives $t(4,p)=2$, and then $g(4,p^e)=g(4,p)$ for all $e\geqq 1$ as in Case 1 of Lemma 1; for $g(4,p)$ use the table above, with the fact that $g(4,p)=2$ for $p>(4-1)^4=81$. If $p\equiv 3 \pmod 4$ then $(4,p-1)=2$, so that $g(4,p)=g(2,p)=2$; for $e>1$, $g(4,p^e)\geqq 3$ by part (2) of Lemma 2, but $g(4,p^e)\leqq 3$ by [1], 2.2. For p satisfying $29<p\equiv 5 \pmod 8$ we have $g(4,p)=2$ from the table, and $g(4,p^e)=3$ for $e>1$ as in the preceding case. This leaves only $p=5$, $p=13$, and $p=29$. For $p=5$, compute $g(4,5)=4$ and $g(4,25)=5$ directly, and conclude $g(4,5^e)=5$ for all $e>1$ by [1], 2.2. The fourth powers mod 13 are $0,1,3,9$ so that $t(4,13)=g(4,13)=3$; then [1], 2.2 gives $g(4,(13)^e)=3$ for all $e>1$. Mod 29, the fourth powers are $0,1,7,16,20,23,24,25$, and $t(4,29)=4$; since $g(4,29)=3$ (from the table), part (3) of Lemma 2 gives $g(4,(29)^e)=4$ for $e>1$. This completes the proof.

Clearly, the same arguments lead to a solution for any even k:

THEOREM 4. *Let* $k=2^a p_1^{e_1}\cdots p_s^{e_s}$ *(p_i distinct odd primes, $e_i\geqq 1$, $a\geqq 1$). Denote by $q_1,\dots,q_t$ the set of all odd primes $q<(k-1)^4$ different from $p_1,\dots,p_s$. Put $g(k,2^e)=b_e$ for $1<e\leqq 2a+1$, put $c_{ij}=g(k,p_i^j)$ for $1\leqq i\leqq s$, $1\leqq j\leqq 2e_i+1$ and put $d_l=g(k,q_l)$ for $1\leqq l\leqq t$. Then the $s+t+2(a+\Sigma e_i)$ numbers $\{b_e,c_{ij},d_l\}$ determine $g(k,n)$ for all n. Explicitly, for any prime p and any $e\geqq 1$ we have*:
Case 1. $p=2$. *Then* $g(k,p^e)=1$ *if* $e=1$, b_e *if* $1<e\leqq 2a+1$, *and* b_{2a+1} *otherwise.*
Case 2. $p=p_i$. *Then* $g(k,p^e)=c_{ij}$ *where* $j=\min(e,2e_i+1)$.
Case 3. $p\neq 2$, $p\neq p_i$ *for all* $1\leqq i\leqq s$, *and* $p<(k-1)^4$. *Then* $p=q_l$ *for some* l, $1\leqq l\leqq t$, *and* $g(k,p^e)=d_l$ *if* $t(k,p)\leqq g(k,p)$; *if* $t(k,p)=g(k,p)+1$ *then* $g(k,p^e)=d_l$ *for* $e=1$ *and* d_l+1 *for* $e>1$.
Case 4. $p>(k-1)^4$, $p\equiv 1 \pmod{2^{a+1}}$. *Then* $g(k,p^e)=2$.
Case 5. $2\neq p>(k-1)^4$, $p\not\equiv 1 \pmod{2^{a+1}}$. *Then* $g(k,p^e)=2$ *for* $e=1$, 3 *for all* $e>1$.

The proof of Theorem 4 is immediate from Lemma 2.

A comparison of Theorem 4 (for even k) and Theorem 2 (for odd k) shows that we need slightly more information to determine $g(k,n)$ when k is even; in particular, for even k we have to know whether or not $t(k,p)\leqq g(k,p)$, for a certain finite set of primes p. We are therefore led to the following

PROBLEM. Let k be even, let p be prime, and assume $p\equiv 1 \pmod k$ and $k+1<p<(k-1)^4$. Find necessary and sufficient conditions, hopefully in the form of congruences relating k and p, for $t(k,p)\leqq g(k,p)$, i.e., for -1 to be a sum of fewer than $g(k,p)$ kth powers mod p.

Lemma 2 (parts 1 and 4) gives a sufficient condition, namely that $p-1$ be slightly "more even" than k, and shows that it is also necessary whenever $g(k,p)=2$, in particular for $p>(k-1)^4$. It is not necessary in general: $g(4,13)=g(4,29)=t(4,13)=3$, but $t(4,29)=4$. This example suggests, too, that the answer may have to do with the distribution of the kth powers mod p; the fourth powers $1,3,9$ mod 13 are "more evenly spaced" than the fourth powers $1,7,16,20,23,24,25$ mod 29.

References

1. C. Small, Waring's problem mod n, this MONTHLY, 84 (1977) 12–25.
2. ———, Powers mod n, Math. Mag., 50 (1977) 12.

DEPARTMENT OF MATHEMATICS, QUEEN'S UNIVERSITY, KINGSTON, ONTARIO, CANADA.

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Department of Mathematics, University of Wisconsin, Madison, WI 53706. (Until July 15, 1977: Laboratoire Calcul des Probabilités, Université de Paris, T.56, 4 Place Jussieu, 75–230 Paris, France.)

PLANES, CUBES AND CENTER-REPRESENTABLE POLYTOPES

LAMBERT S. JOEL, DOUGLAS R. SHIER AND MARJORIE L. STEIN

Abstract. Given an n-dimensional hypercube $C(n)$, is there a hyperplane that meets the interior of every facet of $C(n)$? Contrary to the impression given by the cases $n=1$ and $n=2$, the answer is "yes" for all $n>2$. Furthermore, for $n>3$, the hyperplane can also be chosen to contain any prescribed point of $C(n)$. Such results imply that the center of the smallest hyper-rectangle having sides parallel to the coordinate planes and enclosing a convex polytope K does not necessarily lie in K. A practical estimation problem motivating the inquiry is described.

The following geometrical question arose during the investigation of an estimation problem. Given a closed n-dimensional hypercube $C(n)$ in n-dimensional Euclidean space $\mathbf{E}^n$, is there an $(n-1)$-dimensional hyperplane that meets the interior of every $(n-1)$-dimensional face (or **facet**) of $C(n)$? If $n=1$ or 2, the answer is clearly "no." For example, it is impossible for a straight line to intersect the interior of every edge of a square. For $n>2$ however, the answer is "yes." In fact, when $n>3$ the hyperplane can also be chosen to contain any given point of $C(n)$, even a boundary point! In this note we prove these assertions and relate such results to the practical problem that supplied the original motivation.

1. Result for $n>2$.

THEOREM 1. *Given a closed hypercube $C(n)$ in n-dimensional Euclidean space $\mathbf{E}^n$, there is a hyperplane that meets the interior of every facet of $C(n)$ if $n>2$.*

Proof. Denote a point in $\mathbf{E}^n$ by $X=(x_1,\ldots,x_n)$. Without loss of generality, let

$$C(n)=\{X\in\mathbf{E}^n:0\leqq x_i\leqq 2\quad\text{for}\quad i=1,\ldots,n\}.$$

Define a hyperplane

$$\mathbf{H}=\left\{X\in\mathbf{E}^n:\sum_{i=1}^{n}x_i=n\right\}. \tag{1}$$

It is then straightforward to verify that for $n>2$ the hyperplane $\mathbf{H}$ meets $C(n)$ in the desired manner.

2. Stronger result for $n>3$.

THEOREM 2. *When $n>3$, there are hyperplanes that not only meet the interior of each facet of $C(n)$ but also contain any prescribed point of $C(n)$.*

Proof. Without loss of generality, again define

$$C(n)=\{X\in\mathbf{E}^n:0\leqq x_i\leqq 2\quad\text{for}\quad i=1,\ldots,n\}.$$

Suppose $n>3$ and $Z=(z_1,\ldots,z_n)$ is any given point of $C(n)$. Let $Y=(y_1,\ldots,y_n)$ denote a permutation of components of Z such that the entries are arranged in nonincreasing order,

$$y_1\geqq y_2\geqq\cdots\geqq y_n.$$

By symmetry, there exists a hyperplane of the desired type passing through Z if and only if there exists a hyperplane of the desired type passing through Y. Hence it suffices to consider points Y with components in nonincreasing order.

Define a hyperplane

$$H' = \left\{X \in \mathbf{E}^n : \sum_{i=1}^{n} (-1)^{i-1} x_i = r\right\}, \tag{2}$$

where $r = \sum_{i=1}^{n} (-1)^{i-1} y_i$. Clearly $Y \in \mathbf{H}'$ and $0 \leqq r \leqq 2$. It remains to show $\mathbf{H}'$ meets the interior of every facet of $C(n)$. If $r = 2$ then

$$2 = y_1 - \left[\sum_{i=2}^{n} (-1)^i y_i\right],$$

which (since $y_1 \leqq 2$) implies $y_1 = 2$ and $\sum_{i=2}^{n} (-1)^i y_i = 0$. Again by symmetry, a hyperplane of the desired type passes through $(2, y_2, \ldots, y_n)$ if and only if a hyperplane of the desired type passes through $(0, y_2, \ldots, y_n)$. The components of this latter point can be arranged in nonincreasing order as $(y_2, \ldots, y_n, 0)$, corresponding to a value of $r = \sum_{i=2}^{n} (-1)^i y_i = 0$. Hence it suffices to consider only those points Y such that $0 \leqq r < 2$.

To show $\mathbf{H}'$, as defined in (2) with $0 \leqq r < 2$, meets the interior of every facet of $C(n)$, it is enough to exhibit points of $\mathbf{H}'$ lying on the interior of the facets defined by $x_1 = 0$, $x_2 = 0$, $x_1 = 2$, $x_2 = 2$, because the coordinate labels in (2) are arbitrary. That is, it suffices to test the extreme values 0 and 2 only once for each coordinate appearing in (2) with positive sign and once for each coordinate appearing with negative sign. In what follows, define

$$\delta_n = \begin{cases} 0 & \text{if } n \text{ is even} \\ 1 & \text{if } n \text{ is odd} \end{cases}$$

and choose ε such that $0 < \varepsilon < (2 - r)/2$.

CASE $x_1 = 0$: Let $x_2 = \varepsilon$, $x_3 = r + (2 - \delta_n)\varepsilon$, $x_i = \varepsilon$ for $i > 3$.

CASE $x_2 = 0$: Let $x_1 = (1 + \delta_n)\, r/2 + \varepsilon/2$, $x_3 = (1 - \delta_n)\, r/2 + \varepsilon/2$, $x_4 = (1 + \delta_n)\varepsilon$, $x_i = \varepsilon$ for $i > 4$.

CASE $x_1 = 2$: Let $x_2 = x_4 = (1 + \delta_n)\, \varepsilon/2 + (2 - r)/2$, $x_3 = \varepsilon$, $x_i = \varepsilon$ for $i > 4$.

CASE $x_2 = 2$: Let $x_1 = x_3 = (2 + r)/2 + \varepsilon/2$, $x_4 = (1 + \delta_n)\, \varepsilon$, $x_i = \varepsilon$ for $i > 4$.

By symmetry, the points exhibited above show $\mathbf{H}'$ meets the interior of every facet of $C(n)$. Since $\mathbf{H}'$ also contains the given point Y, the theorem is proved.

3. The case when $n = 3$. The results for the case $n = 3$ are stated below in Theorem 3. The proof, which again involves the straightforward examination of a number of cases, is omitted.

THEOREM 3. *There are precisely* 14 *points of* $C(3)$ *through which no plane meets the interior of each facet of* $C(3)$. *Eight of these points are the vertices of the cube. The remaining six points are the centers of the six facets of the cube.*

4. Center-representable polytopes. The problem motivating the above results was that of estimating the quantities of each of m models of a product sold by each of the k companies in an industry. Each company specializes in producing a known subset of the m models; but because of competition and secrecy in the industry, the number of units a_{ij} of model j sold by company i cannot be directly ascertained. What can be inferred from other data sources, however, are the total number r_i of units of all models sold by *each* company i as well as the total number c_j of units of model j sold by *all* companies.

The desired but unknown quantities a_{ij} form a nonnegative $k \times m$ matrix $A = (a_{ij})$ with certain entries specified to be zero, since a given company may simply not market certain models of the product. Given the row and column totals r_i and c_j, one would like to find "good" estimates for the

entries a_{ij} of A. In general there is no unique set of nonnegative a_{ij}'s that can be determined from the constant row and column sums together with the specified zero entries. In fact, the set of all feasible matrices A forms a convex set in km-dimensional space, since if A_1 and A_2 are feasible then $\lambda A_1 + (1-\lambda) A_2$ is certainly also feasible for $0 \leqq \lambda \leqq 1$. The problem then becomes one of choosing "good" representative points from this closed and bounded convex set.

One seemingly reasonable way of choosing a representative point from a closed and bounded convex set S is to pick the center of the smallest (hyper) rectangle enclosing S and having sides parallel to the coordinate planes. Such an enclosing rectangle will be called the **carton** of S. Its center would have the property that along each dimension the maximum error between the estimated $\hat{a}_{ij}$ and the "true" a_{ij} would be minimized over all possible "true" values for a_{ij}. This fact implies that the center is a minimax location on S with regard to Chebyshev distances. That is, the center c is a point whose maximum distance $d(c, x)$ to any point $x \in S$ is minimum, where $d(c,x) = \max_i |c_i - x_i|$. Unfortunately, this calculated center need not necessarily belong to the set S and may therefore be unsuitable as a "representative point" for S. Theorem 1 will be used to demonstrate such a situation when the convex set is a polytope.

Let us call a closed and bounded set K **center-representable** (CR) if K contains the center of its carton. If $n = 2$, then all closed convex polytopes are CR, as indeed are all closed convex sets in $\mathbf{E}^2$. This follows because if the center of the carton were not in the set, there would be a line through the center separating an entire face of the carton from the convex set. Hence a smaller circumscribing rectangle could have been constructed as the carton.

THEOREM 4. *For $n > 2$ there exist convex polytopes in $\mathbf{E}^n$ that are not CR.*

Proof. Return to the hyperplane defined by (1) and observe that $\mathbf{H}$ contains the center $\mathbf{1} = (1, 1, \ldots, 1)$ of the hypercube $C(n)$. Then a slight perturbation of $\mathbf{H}$ parallel to itself will produce a new hyperplane that still cuts all facets of $C(n)$ but which does not contain $\mathbf{1}$. The new hyperplane divides $C(n)$ into two polytopes and the carton for either polytope is simply $C(n)$ itself. Since one of these polytopes does not contain the center $\mathbf{1}$, we have thus obtained for $n > 2$ a convex polytope that is not CR.

5. Concluding remarks. The notion of center-representability motivated framing Theorems 1, 2, and 3 in terms of hypercubes. These results, however, can be stated with "parallelepiped" replacing "cube" throughout, since any n-dimensional parallelepiped can be mapped into the n-cube $C(n)$ by an affine transformation T of the form

$$x_i' = c_{i0} + \sum_{j=1}^{n} c_{ij} x_j, \quad \text{where } i = 1, \ldots, n.$$

Suppose P is a given point in the parallelepiped; then a hyperplane meeting the interior of each facet of $C(n)$ could be passed through P', the image of P under the transformation T, and the inverse transformation would yield the desired hyperplane. Exceptions in the three-dimensional case will be, as before, the eight vertices as well as the six "centers" of the parallelepiped facets.

Since the polyhedron in the motivating example was found not to be CR, a modification of Deming's method [**1**] was eventually used to obtain a "solution" matrix. When a sample of observed values is available but no zero entries are specified, the estimation problem of Section 4 can be solved using an alternating row and column adjustment method known to converge in such cases [**1**, Ch. 7; **2**]. When zero entries occur, such techniques can break down [**3**].

This work was undertaken while the latter two authors were National Academy of Sciences — National Research Council Postdoctoral Research Associates at the National Bureau of Standards, Washington, D.C. 20234. The authors appreciate the helpful suggestions of the referee.

References

1. W. E. Deming, Statistical Adjustment of Data, Wiley, New York, 1943.
2. C. T. Ireland and S. Kullback, Contingency tables with given marginals, Biometrika, 55 (1968) 179–188.
3. R. Sinkhorn, A relationship between arbitrary positive matrices and doubly stochastic matrices, Ann. Math. Statist., 35 (1964) 876–879.

L. S. Joel and D. R. Shier: Applied Mathematics Division, National Bureau of Standards, Washington, D. C. 20234.
M. L. Stein, U.S. Postal Service, Washington, D.C. 20260.

THE SCHAUDER FIXED POINT THEOREM FOR NONEXPANSIVE MAPPINGS

W. G. Dotson, Jr. and W. R. Mann

It is generally acknowledged that the Schauder fixed point theorem for continuous mappings is rather sophisticated. The purpose of this note is to show that if one restricts attention to nonexpansive mappings ($\| Tx - Ty \| \leqq \| x - y \|$ for all x, y in the domain of T), then this theorem becomes very easy to prove.

Theorem. *Suppose C is a compact convex subset of a Banach space, and T is a nonexpansive self-mapping of C. Then T has a fixed point in C.*

Before proving this theorem, we shall prove a short lemma.

Lemma. *For any fixed $y \in C$ and any fixed positive integer n, define a function G on C by*

$$G(u) = \left(\frac{1}{1+n}\right) y + \left(\frac{n}{1+n}\right) Tu \quad \text{for all} \quad u \in C.$$

Then G has a unique fixed point in C.

Proof. First observe that $G(C) \subset C$ since $y \in C$, $Tu \in C$ $(T(C) \subset C)$ and

$$G(u) = \frac{1}{1+n} y + \frac{n}{1+n} Tu$$

which is a convex combination of y and Tu and is therefore in C since C is convex. Next observe

$$\| G(u_2) - G(u_1) \| = \| y + nTu_2 - y - nTu_1 \| / (1+n) = \frac{n}{n+1} \| Tu_2 - Tu_1 \| \leqq \frac{n}{n+1} \| u_2 - u_1 \|.$$

So G is a contraction on the complete metric space C, and hence has a unique fixed point in C. Q.E.D.

Proof of Theorem. Let y be a fixed element of C, and consider $n = 1, 2, 3, \ldots$. For each such n the corresponding G has a unique fixed point in C which we will denote by x_n. Thus for each n,

$$x_n = G(x_n) = \frac{1}{n+1} y + \frac{n}{n+1} Tx_n.$$

Hence,

$$x_n - Tx_n = \frac{1}{n+1} y + \frac{n}{n+1} Tx_n - \frac{1}{n+1} Tx_n - \frac{n}{n+1} Tx_n = \frac{1}{n+1}(y - Tx_n).$$

Since C is a compact metric space, a subsequence of x_n converges to some $x \in C$ and this x is a fixed point of T. Q.E.D.

Of course, uniqueness cannot be asserted, e.g., consider the reflection of a closed disc about one of its diameters.

DEPARTMENT OF MATHEMATICS, NORTH CAROLINA STATE UNIVERSITY, RALEIGH, NC 27607.
DEPARTMENT OF MATHEMATICS, UNIVERSITY OF NORTH CAROLINA, CHAPEL HILL, NC 27514.

ON MARGINAL DENSITY FUNCTIONS OF CONTINUOUS DENSITIES II

M. J. PELLING AND ALBERT VERBEEK

In [**3**] L. E. Clarke studies the marginal density functions of two-dimensional random variables (X, Y) which have a continuous, finitely valued density function $h(x, y)$. Three nice examples are presented, where such marginals badly fail to be continuous. The following theorem characterizes this class of marginals. Its corollary answers the two questions that conclude [**3**].

DEFINITION. (Cf. Baire [**2**] or Hobson [**4**; p. 307f]). A function $f: R \to (-\infty, +\infty]$ is *lower semi-continuous*, abbreviated *lsc*, if for all x, $\liminf_{x' \to x} f(x') \geqq f(x)$.

THEOREM. *A function f is the marginal density function of some continuous and finitely valued probability density function $h: R^2 \to R$ if and only if*

(i) $f: R \to [0, +\infty]$,

(ii) *f is lsc (and hence Borel measurable), and*

(iii) $\int_{-\infty}^{+\infty} f(x)dx = 1$.

Proof. The proof follows easily from the fact ([**1**] and [**5**; p. 150]) that a function $g: R \to [0, +\infty]$ is lsc if and only if $g = \Sigma_n g_n$ for a series of continuous functions $g_n: R \to [0, +\infty)$. Indeed, given the density $h(x, y)$ we have

$$f(x) = \sum_{n=-\infty}^{+\infty} \int_n^{n+1} h(x, y)dy.$$

Conversely, given f satisfying the conditions of the theorem, let $f = \Sigma_n g_n$ for continuous $g_n: R \to [0, +\infty)$, and let $g^*: R \to [0, +\infty)$ be continuous with support in $[0, 1]$ and $\int_{-\infty}^{+\infty} g^*(x)dx = 1$. Then $f(x)$ is the marginal density function of the continuous density $h(x, y) = \Sigma_n g_n(x)g^*(y + n)$.

COROLLARY. *A marginal density function f, satisfying the conditions of the theorem, is continuous (possibly $+\infty$-valued) at all points of a dense G_δ set.*

Not every probability density function on R is a.e. equal to such a marginal.

The first part of the corollary is a special case of the well-known theorem of Baire that the continuity points of a pointwise limit of a sequence of continuous functions constitute a dense G_δ set. (See [**2**] or [**5**; p. 264f].)

The second part is illustrated and proved by the following example. Let c be the characteristic function of a closed, nowhere dense subset of R of Lebesgue measure 1, e.g., of a suitably chosen Cantor middle third set. Suppose that f is an lsc function, and that c is a.e. equal to f. Then any open interval contains points x such that $c(x) = f(x) = 0$. From the definition above we see that $f(x) \leqq 0$ for all x. Thus c is a probability density function, that is not a.e. equal to any lsc function.

For a survey of lsc functions see [**2**] or [**4**] and [**5**]. A more detailed description of similarities and dissimilarities between continuity, (a.e.) lower semi-continuity and measurability, in the situation discussed above, is given in [**6**].

References

1. R. Baire, Sur les séries à termes continus et tous de même signe, Bull. Soc. Math. France, 32 (1904) 125–128.
2. ———, Leçons sur les fonctions discontinues, Gauthier-Villars, Paris, 1905.
3. L. E. Clarke, On marginal density functions of continuous densities, this MONTHLY, 82 (1975) 845–846.
4. E. W. Hobson, The Theory of Functions of a Real Variable; Vol. I, 3rd ed., Cambridge University Press, 1927, (Dover reprint, New York, 1957).
5. ———, The Theory of Functions of a Real Variable; Vol. II, 2nd ed., Cambridge University Press, 1926, (Dover reprint, New York, 1957).
6. A. Verbeek, Lower semi-continuity of probability density functions, preprint, available from the author.

DEPARTMENT OF MATHEMATICS, P.M.B. 1154, UNIVERSITY OF BENIN, BENIN CITY, NIGERIA.
SCHAEPMANLAAN 9, AMERSFOORT, NETHERLANDS.

A NOTE ON AN EQUATION RELATED TO THE PELL EQUATION

MORRIS NEWMAN

Abstract. It is shown that the diophantine equation $x^2 - dy^2 = -1$ has solutions, provided that $d = p_1 p_2 \cdots p_r$ where r is 2 or odd and $p_1, p_2, \ldots, p_r$ are distinct primes congruent to 1 modulo 4 such that $(p_i/p_j) = -1$, $i \neq j$.

Let $d > 1$, d square-free. The conditions under which the diophantine equation

$$x^2 - dy^2 = -1 \tag{1}$$

possesses solutions are still not fully known, although of course, it is known that the Pell equation

$$x^2 - dy^2 = 1 \tag{2}$$

always has non-trivial solutions. If $d = p$, where p is a prime $\equiv 1$ modulo 4, then (1) has a solution; and if $d = pq$, where p, q are distinct primes $\equiv 1$ modulo 4 such that $(p/q) = -1$, then (1) also has a solution, as was shown by Dirichlet in [**1**]. Dirichlet also treated the case when d is the product of three distinct primes. The theorem in this note is a direct generalization of Dirichlet's results and has apparently not been noticed previously.

We shall prove

THEOREM. *Let r be 2 or odd. Let $p_1, p_2, \ldots, p_r$ be distinct primes such that*

$$p_i \equiv 1 \bmod 4, \quad 1 \leqq i \leqq r, \tag{3}$$

$$(p_i/p_j) = -1, \quad 1 \leqq i, j \leqq r, i \neq j. \tag{4}$$

Put $d = p_1 p_2 \cdots p_r$. Then the diophantine equation (1) *has a solution.*

Proof. Assume that (1) has no solutions. Let (ξ, η) be the fundamental solution of (2). We have from (3) that $\xi^2 \equiv \eta^2 + 1 \bmod 4$, which implies that ξ is odd and η even. Thus we have

$$\frac{\xi - 1}{2d_1} \cdot \frac{\xi + 1}{2d_2} = \frac{\eta^2}{4},$$

where $d_1 d_2 = d$, and $((\xi - 1)/2d_1, (\xi + 1)/2d_2) = 1$. It follows that

$$\frac{\xi - 1}{2d_1} = u^2, \quad \frac{\xi + 1}{2d_2} = v^2, \quad d_2 v^2 - d_1 u^2 = 1.$$

We note that $u \neq 0$, since $\xi \neq 1$.

Suppose now that neither d_1 nor d_2 is 1. Since r is either 2 or odd, either d_1 or d_2 must be the product of an odd number of primes. Suppose that d_1 is the product of an odd number of primes, and let p be any prime dividing d_2. Because of (4) we have that $(d_1/p) = -1$, which contradicts the fact

that $d_1u^2 \equiv -1 \bmod p$. Similarly, the case when d_2 is the product of an odd number of primes may also be shown to be impossible. It follows that $d_1 = 1$, $d_2 = d$, or $d_1 = d$, $d_2 = 1$. But $d_1 = 1$, $d_2 = d$ gives $u^2 - dv^2 = -1$, which contradicts the assumption that (1) has no solutions. Thus $d_1 = d$, $d_2 = 1$, and so $v^2 - du^2 = 1$. We have

$$u^2 = \frac{\xi - 1}{2d}, \qquad v^2 = \frac{\xi + 1}{2}, \qquad u^2v^2 = \frac{\xi^2 - 1}{4d} = \frac{\eta^2}{4}.$$

Hence $uv = \eta/2$, and so $u < \eta$. But this contradicts the fact that (ξ, η) is the fundamental solution of (2). Hence the assumption that (1) has no solutions is false, and the theorem is proved.

Reference

1. G.L. Dirichlet, Abh. Akad. Wiss. Berlin, (1834) 649–664, Werke 1, 219–236.

NATIONAL BUREAU OF STANDARDS, WASHINGTON, DC 20234.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics and Statistics, The University of Calgary, Calgary, Alberta, Canada, T2N 1N4. (From July 1976 to June 1977: Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, 16 Mill Lane, Cambridge CB2 1SB, England.)

WHEN IS THE GRAPH OF A TRIANGULATION UNIQUELY 4-COLORABLE?

PETER KLEINSCHMIDT

A graph is said to have an ***n*-coloring** if there is a partition of its vertices into n classes, called the **color classes,** in such a way that no adjacent vertices belong to the same color class. A graph with exactly one n-coloring is called **uniquely *n*-colorable.** Uniquely colorable graphs have been studied by Harary, Hedetniemi and Robinson in **[3]** and by Chartrand and Geller in **[2]**. It is well known that the 4-color-conjecture is equivalent to the conjecture that the graph of every triangulation of the plane possesses a 4-coloring.

One may ask which graphs are 'extreme' with respect to this conjecture, i.e., graphs of triangulations which are uniquely 4-colorable.

We consider a special class of such graphs, the edge-graphs of stack polytopes (see **[1]**, **[4]** and **[5]**). A 3-dimensional polytope is called a **stack polytope** if its boundary complex is isomorphic to a complex obtained from the boundary complex of a tetrahedron by successively applying barycentric subdivisions to triangles. We can now formulate a conjecture for uniquely 4-colorable graphs.

CONJECTURE. *The graph of a triangulation of the plane is uniquely* 4-*colorable if and only if it is combinatorially isomorphic to the graph of a stack polytope.*

It is obvious that the graph of a stack polytope is uniquely 4-colorable. To show the more interesting part of the conjecture, one would have to prove that no 4-connected planar graph, which

triangulates the plane, possesses exactly one coloring. From this it would follow that a uniquely 4-colorable graph G of a triangulation consists of two components G_1 and G_2 which have a circuit of length 3 in common (see [4]). By induction on the number of vertices it would follow that G_1 and G_2 and therefore G are graphs of stack-polytopes.

References

1. A. Altshuler, Manifolds in stacked 4-polytopes, J. Combinatorial Theory, 10 (1971) 198–239.
2. G. Chartrand and D. P. Geller, On uniquely colorable planar graphs, J. Combinatorial Theory, 6 (1969) 271–278.
3. F. Harary, S. T. Hedetniemi and R. W. Robinson, Uniquely colorable graphs, J. Combinatorial Theory, 6 (1969) 264–270.
4. P. Kleinschmidt, Eine graphentheoretische Kennzeichnung der Stapelpolytope, Archiv der Mathematik, to appear.
5. G. C. Shephard, Subpolytopes of stack polytopes, Israel J. Math., 19 (1974) 292–296.

Mathematisches Institut, Ruhr Universität Bochum, D 4630 Bochum, Federal Republic of Germany.

CLASSROOM NOTES

Edited by Richard A. Brualdi

Material for this Department should be sent to Richard A. Brualdi, Department of Mathematics, University of Wisconsin, Madison, WI 53706. (Until July 15, 1977: Laboratoire Calcul des Probabilités, Université de Paris, T.56, 4 Place Jussieu, 75–230 Paris, France.)

KLEIN'S 'ERLANGER PROGRAMM' AND THE GEOMETRY OF AN INFINITESIMAL REGION

Cyril W. L. Garner

The ideals of Felix Klein's *Erlanger Programm* are used to provide a simple proof of the well-known result that the geometry of an infinitesimal region is euclidean.

It is well known that the geometry of an infinitesimal region of the hyperbolic plane is euclidean. The familiar proofs use limiting arguments applied to the metric of the hyperbolic plane ([1], pp. 211–212) or to hyperbolic trigonometry ([1], p. 238). These proofs, while not difficult, require a fair amount of background knowledge. The following is a very simple proof of this result, and is an interesting application of Felix Klein's famous *Erlanger Programm* of 1872.

According to the *Erlanger Programm* (see for example [4]), a geometry is any theory that studies the action of a group of transformations on a set. For example, the classical inversive plane is obtained from the classical euclidean plane by adjoining an ideal point which is both the centre and a point of every euclidean straight line, regarded as circles of infinite radius ([2], p. 83 or [5], p. 205). Thus the classical inversive plane has as "points" the euclidean points and this ideal point, as "M-circles" the euclidean circles and straight lines, and as "incidence" the usual euclidean incidences plus the above new incidences used in adjoining the ideal point. Inversive geometry is then the study of the group $\mathscr{C}$ of circular transformations ([5], pp. 208–212), and euclidean geometry is the study of a subgroup of $\mathscr{C}$ which stabilizes a point.

Analogously, hyperbolic geometry is the study of a subgroup of $\mathscr{C}$ which stabilizes an M-circle, A. Since this M-circle can be either a euclidean circle or straight line, this concept gives rise to both the

conformal model of the hyperbolic plane (due to Poincaré) and the half-plane model, which are of course isomorphic. Both models were used extensively by Klein and Fricke — see **[6]** for an illuminating discussion of the fascinating diagrams appearing in **[3]**.

In the conformal model of the hyperbolic plane, "hyperbolic points" are those points interior to A, and "hyperbolic lines" are the intersections of the interior of A with those M-circles which are orthogonal to A, while "incidence" is obvious. Alternatively, we could replace the interior of A by the exterior of A. If A is a euclidean straight line, then the interior and exterior of A are simply the two half-planes determined by A in the euclidean plane.

Consider the M-circle A to be a euclidean circle. The question naturally arises — what happens if the radius of A is arbitrarily small? Here it is helpful to think of hyperbolic geometry as the study of either the *interior* or of the *exterior* of A — they are isomorphic. With the first viewpoint, as the radius of A approaches 0, we are looking at the geometry of an infinitesimal region. But with the second viewpoint, as the radius of A approaches 0, we are studying the properties of sets in the plane under a subgroup of $\mathscr{C}$ which stabilizes a point. This is euclidean geometry, and so we have a new proof of the result that the geometry of an infinitesimal region is always euclidean.

References

1. H. S. M. Coxeter, Non-euclidean Geometry (third edition), University of Toronto Press, Toronto, 1957.
2. ———, Introduction to Geometry (second edition), John Wiley, New York, 1969.
3. R. Fricke and F. Klein, Vorlesungen über die Theorie der automorphen Funktionen Vol. 1, Teubner, Leipzig (1897). Reprinted by Johnson Reprint, New York, 1965.
4. C. W. L. Garner, Felix Klein's 'Erlanger Programm', Carleton Mathematical Series No. 101, Ottawa, 1974.
5. H. W. Guggenheimer, Plane Geometry and Its Groups, Holden-Day, San Francisco, 1967.
6. W. Magnus, Noneuclidean Tesselations and Their Groups, Academic Press, New York, 1974.

DEPARTMENT OF MATHEMATICS, CARLETON UNIVERSITY, OTTAWA, CANADA K1S 5B6.

ON THE CONTINUOUS DEPENDENCE OF THE ROOTS OF A POLYNOMIAL ON ITS COEFFICIENTS

D. J. UHERKA AND ANN M. SERGOTT

One hears frequent reference to the facts that the (complex) roots of a polynomial depend continuously on the coefficients of the polynomial and that the eigenvalues of a matrix depend continuously on the entries of the matrix. However, the proofs of these equivalent results do not seem to appear in the standard texts in advanced calculus, complex variables, numerical analysis, matrix theory, etc., where these theorems should be of interest at least as exercises.

There have been many proofs, in the literature, of the above theorem for polynomials, with and without taking into account multiplicities of roots. See, for example, **[1]**–**[6]** for proofs, some of which are relatively short, elementary, and ingenious. The use of Rouché's theorem for polynomials **[4]**, makes the proof of the continuity of the roots very easy and straightforward. However, if a student were simply given the task of proving the continuity of the roots of a polynomial, it would take a good deal of ingenuity for him to discover Rouché's theorem, not having seen it before. The other proofs that appear in the references also seem to require a certain degree of ingenuity.

It seemed to us that one should be able to find a proof which is even shorter and more elementary than those that have appeared, and which requires little ingenuity and would be appropriate for undergraduate courses of the type listed above. We believe the proof given below satisfies these requirements. The theorem we have in mind is the one that, stated simply, says: Given $\varepsilon > 0$, if the coefficients of a polynomial are changed by a sufficiently small amount, then the roots (including

multiple roots) of the polynomial change by no more than ε. Proving this result does not require that the n roots of an nth degree polynomial be established as n continuous functions of the coefficients, i.e., we need not prove an implicit function theorem. However, it does require the knowledge that the polynomial has roots. We will assume the fundamental theorem of algebra and the uniqueness of the factorization of a polynomial (which follows easily from the fundamental theorem of algebra, using only elementary algebra). We will also assume the standard theorems about convergence of sequences in complex n dimensional Euclidean space.

Consider the following two nth degree polynomials, one having roots λ_i and the other having roots λ_i^*:

$$p(x) = x^n + a_1x^{n-1} + \cdots + a_n = (x-\lambda_1)\cdots(x-\lambda_n)$$

$$p^*(x) = x^n + a_1^*x^{n-1} + \cdots + a_n^* = (x-\lambda_1^*)\cdots(x-\lambda_n^*).$$

We shall use the following three facts which are immediate consequences of elementary algebra and limit theorems:

(1) If x is a root of $p(x)$, then $|x| \leqq \max\{1, \Sigma_{i=1}^n |a_i|\}$.

(2) If a_i approaches a_i^* $(i = 1,\ldots,n)$ then $p(x) \to p^*(x)$ for all (real or complex) x.

(3) If λ_i approaches λ_i^* $(i = 1,\ldots,n)$ then $p(x) \to p^*(x)$ for all x.

Statements (2) and (3) could be made more precise with ε's and δ's, but the meaning is clear. We are now ready to state and prove the classical root continuity theorem. Since $p(x)$ has n roots, the words "at least" in the theorem below can clearly be replaced by "exactly."

THEOREM. *Suppose λ^* is a root of p^* of multiplicity m and $\varepsilon > 0$. Then for $|a_i - a_i^*|$ sufficiently small $(i = 1,\ldots,n)$, p has at least m roots within ε of λ^*.*

Proof. Suppose not. Then there exists a sequence $\{p_k\}$ of polynomials whose coefficients converge to those of p^*, such that p_k has fewer than m roots within ε of λ^*. Since the coefficients of $\{p_k\}$ converge, they are bounded. Hence, if

$$p_k(x) = (x - \lambda_1^{(k)})\cdots(x-\lambda_n^{(k)}),$$

then by (1), the root sequence $\{(\lambda_1^{(k)},\ldots,\lambda_n^{(k)})\}$ is bounded (in n-dimensional complex space with Euclidean distance) and has a subsequence convergent to some point $(\mu_1,\ldots,\mu_n)$ with at most $m-1$ of the μ_i being equal to λ^*.

If the corresponding subsequence of polynomials is $\{p_{k_i}\}$, then by (2), $p_{k_i}(x) \to p^*(x)$ for all x. But by (3), $p_{k_i}(x) \to (x-\mu_1)\cdots(x-\mu_n)$ which cannot be $p^*(x)$ since at most $m-1$ of the μ_i equal λ^*. This contradiction finishes the proof.

There are examples in some texts which may lead some students to feel that the roots of a polynomial do not change continuously with the coefficients. Wilkinson [7], for instance, gives the example (which also appears in several numerical analysis texts) $p(x) = (x-1)(x-2)\cdots(x-20)$, for which changing the coefficient of x^{19} by only 2^{-23} causes some of the roots to move a distance of about 2 or 3 in the complex plane. This situation, however, should not be interpreted as indicating a discontinuity any more than does the simple function $f(x) = 3\cdot 2^{23}x$, which exhibits a similar behavior when x is changed by 2^{-23}.

Finally, it should be noted that the *real* roots of a polynomial may not depend continuously on the coefficients, as is shown by the example $p(x) = (x^2+2x+c_1)(x^2-4x+c_2)$ in a (real or complex) neighborhood of $c_1 = 1$, $c_2 = 4$.

The authors wish to thank the referee for some helpful comments and suggestions.

References

1. R. C. Gunning and H. Rossi, Analytic Functions of Several Complex Variables, Prentice-Hall, Englewood Cliffs, N.J., 1965, p. 13.

2. R. Iglisch, Zur Stetigkeit der Wurzeln einer algebraischen Gleichung, Jber. Deutsch. Math.-Verein, 7 (1944) 520–521.

3. H. Kneser, Zur Stetigkeit der Wurzeln einer algebraischen Gleichung, Math. Z., 48 (1942) 101–104.

4. M. Marden, The Geometry of the Zeros of a Polynomial in a Complex Variable, Amer. Math. Soc., (1949) 1–4.

5. W. Rudin, Functional Analysis, McGraw-Hill, New York, 1973, p. 239.

6. B. L. van der Waerden, Einführung in die algebraische Geometrie, American edition by Dover, New York, 1945, p. 48.

7. J. H. Wilkinson, Rounding Errors in Algebraic Processes, Prentice-Hall, Englewood Cliffs, N.J., 1963, p. 43.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF NORTH DAKOTA, GRAND FORKS, ND 58202.

THE CURIOUS SUBSTITUTION $Z = \tan \theta/2$ AND THE PYTHAGOREAN THEOREM

ALAN H. SCHOENFELD

In their discussion of the substitution $Z = \tan \theta/2$, calculus texts generally present the substitution and verify that it transforms any rational function of $\sin \theta$ and $\cos \theta$ to a rational function of Z. This approach deprives students of a means for deriving the substitution and a frame of reference for it. The discussion in Section 1, written for students, provides both.* In Section 2 we discuss the form of all substitutions which express $\sin \theta$ and $\cos \theta$ as rational functions of Z.

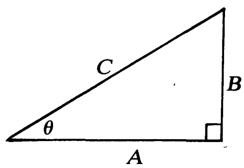


FIG. 1

1. A Derivation. The problem is to find a substitution which expresses $\sin \theta$ and $\cos \theta$ as rational functions of Z. For motivation we examine Figure 1, which provides us with the relationships $\sin \theta = B/C$, $\cos \theta = A/C$, and $C^2 = A^2 + B^2$. Now if we could find *polynomials* $A(Z)$, $B(Z)$, and $C(Z)$ such that the values of $A(Z)$, $B(Z)$, and $C(Z)$ always formed a right triangle (Figure 2), then $\sin \theta = B(Z)/C(Z)$ and $\cos \theta = A(Z)/C(Z)$ would be rational functions of Z. By the Pythagorean theorem, a necessary and sufficient condition for this is

$$[C(Z)]^2 = [A(Z)]^2 + [B(Z)]^2. \tag{1}$$

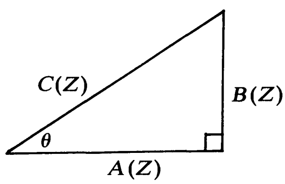


FIG. 2

Thus if we can find three polynomials satisfying (1), we can use Figure 2 to obtain the desired

*This question arose in a conversation with Professor F. Reif, who thought that such a substitution should "make sense" to students.

the form

$$\sin\theta = \frac{2FG}{F^2+G^2}, \qquad \cos\theta = \frac{F^2-G^2}{F^2+G^2},$$

where F and G are polynomials in Z with no common factors. (The roles of $\sin\theta$ *and* $\cos\theta$ *may be reversed, with correction for a minus sign.)*

Proof. As noted in Section 1, we need polynomials A, B, and C in Z such that $C^2 = A^2 + B^2$, or

$$B^2 = (C+A)(C-A).$$

We may assume without loss of generality that A, B, and C have no common factors. Now if $f(Z)$ is any irreducible factor of B, then f cannot divide both $(C+A)$ and $(C-A)$, for then it would divide $A = \frac{1}{2}[(C+A)-(C-A)]$ as well. It follows that $(C+A) = F^2$ and $(C-A) = G^2$, where $B = FG$ and F and G are polynomials in Z with no common factors. Then $A = \frac{1}{2}(F^2 - G^2)$ and $C = \frac{1}{2}(F^2 + G^2)$, and labeling the triangle as in Figure 2 provides the desired values of $\sin\theta$ and $\cos\theta$.

GROUP IN SCIENCE AND MATHEMATICS EDUCATION, PHYSICS DEPARTMENT, UNIVERSITY OF CALIFORNIA, BERKELEY, CA 94720.

EMBEDDING GRAPHS IN EUCLIDEAN 3-SPACE

A. K. DEWDNEY

In courses on Graph Theory, especially in the study of embeddings of graphs in various surfaces, it is frequently mentioned, but rarely proved, that any graph can be embedded in R^3. Topological proofs of this fact, such as [**3**, p. 26], involve notions which are outside the scope of most graph theory courses. A more accessible proof is presented in very general form in Grünbaum's book [**2**, p. 62] and, more recently, outlined in White's book [**4**, p. 51]. This proof uses only elementary algebra. First a *moment curve* [**2**] in R^3 is defined by the equation $f(t) = (t, t^2, t^3)$. Next, it is observed that the *Vandermonde determinant,* defined in [**1**, p. 303] as the determinant of the matrix $A = (a_{ij})$, where $a_{ij} = t_i^{j-1}$, may be written $|A| = \Pi_{i>j}(t_i - t_j)$. This observation can even be made a special exercise for students wishing to 'brush up' the algebra. The result can now be given.

THEOREM. *Every graph can be embedded in* R^3.

Proof. Let G be an arbitrary graph with vertices v_k $(k = 1, 2, \ldots, n)$. For each k, let u_k be the point $f(k) = (k, k^2, k^3)$ in R^3. Join u_i and u_j by a straight line segment in R^3 whenever $\{v_i, v_j\}$ is an edge of G. Obviously, the subset H of R^3 so obtained will be an embedding of G provided that no two of the straight line segments meet at an interior point. If $\{u_i, u_j\}$ and $\{u_k, u_l\}$ are two such segments, they can only meet at an interior point if i, j, k and l are all distinct. Moreover, they can only meet at an interior point if the four points $f(i), f(j), f(k), f(l)$ on the moment curve are coplanar, i.e., they all lie on some plane, say $ax + by + cz + d = 0$, where not all of a, b, c, d can be zero. The corresponding system of four linear equations in the unknowns a, b, c, d has a non-trivial solution if and only if the determinant of the coefficient matrix is zero. This determinant is the Vandermonde determinant, and, as observed above,

$$\begin{vmatrix} 1 & i & i^2 & i^3 \\ 1 & j & j^2 & j^3 \\ 1 & k & k^2 & k^3 \\ 1 & l & l^2 & l^3 \end{vmatrix} = (i-j)(i-k)(i-l)(j-k)(j-l)(k-l).$$

The expression on the right cannot be zero, however, since i, j, k, l are all distinct. It follows that $f(i), f(j), f(k), f(l)$ are not coplanar and the theorem is proved.

In Figure 1 below is shown a 3-dimensional plot of the moment curve $f(t) = (t, t^2, t^3)$, $0 \leqq t \leqq 12$. The divisions of x, y and z axes refer to units, tens and hundreds, respectively. For ready visualization, the curve has been projected on both the xy- and xz- axes.

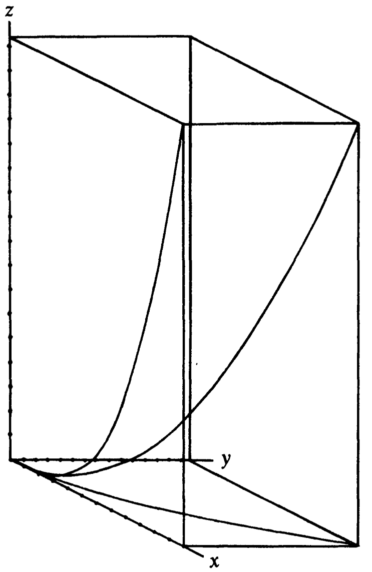


FIG. 1. The moment curve.

This figure may help give students a feeling for why the moment curve works. Of course, there is nothing magical about the curve we have used; many other curves with the "non-coplanarity property" may be found. This raises the interesting problem of characterizing such curves.

References

1. G. Birkhoff and S. MacLane, A Survey of Modern Algebra, Macmillan, New York, 1953.
2. B. Grünbaum, Convex Polytopes, Interscience/Wiley, New York, 1967.
3. B. Rushing, Topological Embeddings, Academic Press, New York, 1973.
4. A. T. White, Graphs, Groups and Surfaces, North Holland, Amsterdam, 1973.

DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF WESTERN ONTARIO, LONDON, ONTARIO, CANADA, M6A 5B9.

EDITED BY PAUL T. MIELKE AND SHIRLEY HILL

Material for this Department should be sent to Paul T. Mielke, Department of Mathematics, Wabash College, Crawfordsville, IN 47933, or to Shirley Hill, Department of Mathematics, University of Missouri, Kansas City, MO 64110.

THE MATHEMATICAL EDUCATION OF WOMEN

MARY GRAY

Since the Carnegie Commission report [6] focused on the lack of mathematical training as a hindrance to equal opportunity for women, increased attention has been devoted to the problem of how to improve the mathematical education of women at all levels. As part of its experimental projects division the National Science Foundation has funded a number of projects concentrating on getting more women into scientific careers, three of the grants for 1974–75 going to studies on women and mathematics. Other foundations have funded, or have expressed an interest in funding, similar projects.

The Johns Hopkins project on intellectually gifted children has devoted some of its efforts to trying to discover why few adolescent girls show, by the project's measurements, precocious mathematical ability. Lynn Fox [3] has tried to foster mathematical talent in junior high girls by various means.

At the other end, in a sense, of the mathematical training of women, Ravenna Helson [4] has studied the characteristics of creative women mathematicians. That creative mathematicians of either sex may be unsocial and self-centered as well as having great power of concentration and terrific tenacity comes as no surprise, but no one suggests that training women in these traits should be undertaken to produce more women mathematicians. Many writers (e.g., Simone de Beauvoir) have, however, pointed out how easy society makes it for women to lack concentration and tenacity.

The problem is not that of attempting to produce female Fields medalists. Such talent is so rare and possibly so little subject to nurture that discussion of how to promote it seems futile. However, the difficulties in the mathematical education of women are apparent at every level. John Ernest [2] noted that in grades 2–12 there is very little difference in the positive attitudes girls and boys express for mathematics, but there is a significant difference in the perception by teachers of how well girls and boys do in the subject. That is, in spite of the fact that girls' achievement in mathematics compares favorably with boys' throughout elementary school, a significantly larger percentage of teachers believed that boys do better in mathematics than girls than believed that girls surpass boys.

The Johns Hopkins study conjectured that the socialization process is so far advanced by junior high age that girls shy away from the study of mathematics and certainly from the intensive devotion to it observed in many of their male contemporaries.

Lucy Sells [7] studied the mathematical preparation of freshmen entering the University of California at Berkeley and found that only 8% of the women (as compared with 57% of the men) had an adequate high school mathematics background for any of the calculus sequences. Thus 92% were effectively foreclosed from careers in mathematical, physical and biological sciences, most social sciences and engineering, all of which at Berkeley require calculus.

Elementary and secondary education. The Sells, Fox and Ernest studies might engender a feeling among college faculty that most of the damage has been done before women reach college. Although it is true that the socialization of girls is a broad societal problem which is not amenable to quick and easy solution, as members of society at large if not specifically as mathematicians, we all have responsibility for seeing that each child has all possible opportunity for realizing her or his potential.

Much of the adverse early childhood conditioning is not specific to mathematics; it is simply the

general female role versus the male role. But there are special problems having to do with the image of mathematics as a man's field — for example, the mother's helping with all the homework except the math and the father's helping with that. And there are still teachers and parents who believe — and lead their students to believe — that girls have more ability for languages while boys have more ability in mathematics.

However, since college and university mathematics faculty are responsible in many institutions for training future elementary and secondary school teachers, they have even more specific responsibilities. If mathematics courses for education majors are the responsibility of mathematics faculty, they are all too frequently looked on as a distasteful part of the curriculum to be dispensed with by as little expenditure of effort as possible. This is particularly unfortunate, for unimaginative and boring teaching can reinforce the dislike and fear that many students bring to these courses. Reform in many cases can begin by the recognition that the fear and dislike exist, and along with them an inadequacy in basic mathematical skills. In fact it is this inadequacy, along with other factors, which breeds the fear and dislike. Some attention needs to be devoted to replacing the negative attitude with a positive one as well as to communicating mathematical knowledge; in particular, the teachers and the students need to feel that the students *can* handle mathematics.

One technique that seems to be effective is to organize a course for teachers around the nature of mathematics and its applications to real life and to supplement regular classroom lectures with workshops or laboratories. The aims are an understanding of the power and beauty of the subject together with the skills to show its usefulness.

The laboratory and workshop methods seem to be very useful for students who come to a course with varied backgrounds and who may be too inhibited to ask their basic questions in a classroom setting. Skillful and experienced teachers with a good feel for mathematics, and for the difficulties people have with it, should be assigned to these courses.

It is not by chance that most education majors, especially in elementary education, are women. Teaching is, of course, a traditionally acceptable occupation for females for the many reasons we need not repeat here. But more important, these are often the women who have shut themselves off from many careers by their attitude toward mathematics, the very attitude that should not be passed on to their students (female or male). Some of them are not in elementary education because they want to be teachers, but because they believe it is a major which allows them to avoid mathematics courses with "real" content. Thus it is all the more important that an attempt be made to change these attitudes toward mathematics.

In addition to improving teacher education, mathematicians can encourage girls to remain in high school mathematics courses long enough to prepare them for college courses and the subsequent career options. Visiting high schools, working with guidance counsellors, examining and revising the recruiting material used by their own institutions, developing early admissions programs — these are but a few of the ways to attack the problem.

Undergraduate mathematics courses. In some institutions many women are effectively closed out of particular career options because they are not prepared for calculus courses and no pre-calculus courses are available. Such courses should be provided not only for women but for others who may have inadequate backgrounds, rather than closing off options for the mathematically disadvantaged. As in the case of courses for education majors, however, the pre-calculus courses are frequently stepchildren of the mathematics department.

It should be remarked that the purpose of remedial work for women students is not necessarily to turn them into mathematics majors, but rather to open up to them options not only in mathematical sciences, but in natural sciences, the quantitative aspects of social sciences, and business and management.

It may be that women students need special encouragement — to seek careers, to speak up in class, to sign up for honor sections. It may be that instructors need to examine the remarks they make, the

preconceived notions they have, the textbooks they use. While most authors have dropped such examples as those showing a negative correlation between bust measurements and IQ, the lawyer "he," the secretary "she," and other sexist references still abound.

Role models. One of the issues which recurs in discussions on the education of women is the existence or non-existence of role models. Many believe that women will be inspired to be mathematicians only by seeing successful women mathematicians. This notion was rejected by most women mathematicians responding to a questionnaire of the AMS Committee on Women; however, since those who dropped out because they had no role models (or for any other reason) were not reached by this survey and those surveyed had obviously managed without, the result is not conclusive.

Surveys of successful research doctorate women have showed that a disproportionately large number held undergraduate degrees from women's colleges [**8, 9**]. In these schools, especially in the past, a large share of the faculty were women and were indeed role models. There may, of course, be other factors accounting for the large percentage of women's college graduates in this group. The most likely reality is that the absence of role models may be a hindrance to some and the presence may be a help to some.

A parallel issue is whether women need to be taught by women because only women understand women's special problems. The analogous question with minorities has also been widely raised. Again, some women students may respond better to women teachers and advisers whereas others may not. What is clearly needed is more sensitivity to the problems of women students on the part of both women and men faculty.

For example, women students are in need of someone either to demonstrate or reassure them that they need not necessarily choose between a career and a family. They need someone who understands the choices that do need to be made, the conflicting claims on their attention, the ease with which women are allowed to quit and the extra encouragement they may thus require in order to continue. Contrary to the impressions of some, there has been virtually no increase in recent years in the number or percentage of women on mathematics faculties, although the percentage of Ph.D.'s in mathematics going to women has increased substantially [**1**]. Thus there are women available to be hired.

In some large institutions, many of the elementary courses are taught by television. Using women faculty and graduate students on the tapes for these is an effective way to reach large numbers of students.

If an institution has no positions open, it can still make some effort to provide role models by inviting women as visiting faculty or as colloquium speakers; there are, after all, many women mathematicians around. For talks particularly for undergraduates, the MAA Visiting Lecturer program has women among its available speakers; for high schools, many sectional MAA Secondary School Lecturer programs list women; under an IBM grant there is a special high school lecturer program directed to girls in three metropolitan areas; more generally, the Association for Women in Mathematics has a speakers' bureau.

Careers. Many question the wisdom of encouraging anyone to seek a career in mathematics in view of today's job market. On the other hand, some contend that women have as much right to be unemployed mathematicians as men do. However, it is clear that students of either sex should be encouraged to realize their potential and to do the best job they can in whatever position they choose at whatever level of mathematical sophistication. Special attention does need to be given to women to keep them from dropping out so soon that they close off paths to careers.

While we are focusing on mathematics as training for careers, we cannot afford to lose sight of the other reasons for studying mathematics. An article in this MONTHLY in 1917 [**5**] addressed itself to the subject of whether women needed higher mathematics; the line of argument was not directed toward career preparation. Many of the reasons for studying mathematics — for general cultural background,

as part of training in logic and organization, have not changed, only intensified; for today anyone who does not know some mathematics, some statistics, some computer science is culturally deprived and ill-equipped to funcion in modern society and to understand its operation.

Another feature of career counselling that does not apply to women only is the necessity of making clear, for example, the many avenues open to one with a bachelor's degree in mathematics; this is probably especially acute for women who see only the teaching role for themselves and not, say, a management career.

Finally, if women are to put to use the mathematical training we try to assure for them, we need non-discriminatory employment practices, again a responsibility which we share not only as mathematicians but as citizens.

The mathematical community's self-interest. It will be to the advantage not only of upcoming generations of women if we encourage and improve the mathematical education of women. If the dropout rate for women at each stage of their study of mathematics were the same as that for men there would be more students to teach. If we combine that with encouraging students of both sexes to go on for advanced degrees only if they are committed to mathematics, the employment picture will ease.

Moreover, with the declining birthrate we all realize that lifetime education is the goal of the future. On the continuing education scene today there are many programs designed especially for women; but hardly any have the mathematical sciences component which could be so useful. This is a disservice to the women and to the mathematics departments who could benefit from the maturity, commitment and talent of these students.

References

1. Annual AMS Survey, Notices Amer. Math. Soc., 21 (1974) 255–258; 22(1975) 303–308.

2. John Ernest, Mathematics and sex, this MONTHLY, 83 (1976) 595–614.

3. Lynn Fox, "Sex difference in mathematical precocity — bridging the gap," in D. P. Keating (ed.), *Intellectual Talent: Research and Development*, The Johns Hopkins Press, Baltimore, 1975.

4. Ravenna Helson, Women mathematicians and the creative personality, J. of consulting and clinical psychology, 36 (1971) 210–220.

5. Emilie N. Martin, Relating to required mathematics for women students, this MONTHLY, 24 (1917) 394–398.

6. Opportunities for Women in Higher Education, Carnegie Commission on Higher Education, McGraw-Hill, New York, 1973.

7. Lucy Sells, Sex and discipline differences in doctoral attrition, Ph.D. Thesis (1975), University of California, Berkeley.

8. M. E. Tidball, Educ. Record, 54, 130 (1973).

9. ———, Baccalaureate origins of women with research doctorates, (to appear)

DEPARTMENT OF MATHEMATICS AND STATISTICS, THE AMERICAN UNIVERSITY, WASHINGTON DC 20016.

MATHEMATICS FOR ELEMENTARY TEACHING: A SMALL-GROUP LABORATORY APPROACH

JULIAN WEISSGLASS

The mathematics departments of many colleges and universities offer courses designed especially for prospective elementary school teachers. The curriculum usually includes properties of sets, properties of the real number system and its subsystems (the positive integers, the integers and the rational numbers), and topics from geometry, algebra and (sometimes) probability. It should be emphasized that these are usually courses in mathematical content and not in the procedures of teaching mathematics.

The University of California, Santa Barbara mathematics department offers a two-quarter course as described above. During the 1973–74 academic year I developed and used curricular materials which enabled me to teach the course in a mathematics laboratory using a small-group approach. An Innovative Projects in University Instruction Grant from the University of California was used to purchase mathematical laboratory equipment from commercial suppliers.

The development of the curricular material and the evaluation of the effects of this approach continue. This article is a progress report. In the succeeding sections I describe: (1) my objectives for the course and the challenge to effective teaching, (2) my assumptions about learning on which the design of the course is based, (3) a description of the laboratory approach, and (4) an evaluation of the project.

1. Objectives and challenges. My objectives for any course in mathematics for prospective elementary school teachers are:

(a) That the student adequately understand the mathematics that is being taught in elementary schools;

(b) That the student develop an awareness of and minimum ability in mathematical reasoning;

(c) That the student appreciate that mathematics (even elementary mathematics) can be interesting, exciting and enjoyable;

(d) That the student with superior mathematical background and abilities be sufficiently challenged so as to make the course interesting;

(e) That the student have the opportunity to discuss mathematical concepts and problems with her/his peers; and

(f) That the student be aware of the relationship of the curriculum to the mathematics he or she will be teaching.

During the time I have taught mathematics to prospective elementary school teachers I have found the achievement of these objectives to be quite challenging. Conversations with colleagues indicated that my experience was not unique. This offered me some relief (at least it wasn't my fault) but little real comfort (the problems were still there). What makes such courses especially challenging to teach?

First, the course is required for all prospective elementary credential applicants. Many of these individuals have had unsuccessful or indifferent experiences with mathematics. The present structure (lectures) and curricular material (textbooks) make it very difficult to achieve the motivation necessary to overcome the effects of these past experiences.

Second, the students enter the class with a wide range of mathematical backgrounds and abilities. A typical class will contain students who have taken only two years of high school mathematics and others who have had one or two years of calculus. Testing indicates a wide range (3rd percentile to 99th percentile) in their knowledge of mathematical concepts and, although I have no supportive data, I am confident that the variation in mathematical aptitude is also very large.

Third, the students express continual dissatisfaction that they are not being taught how to teach mathematics to children. Even after repeated explanations that this is a course in mathematical concepts and not in the procedures of teaching of mathematics, they express discontent at not seeing the relevance of what they are studying to their future careers. This discontent interferes with the learning process.

Fourth, the lecture method, especially if the class is large, hinders any meaningful student participation in the classroom. Mathematicians often state that the only way you learn mathematics is by doing mathematics. If we believe this, we must seriously consider doing something other than lecturing at students.

Finally, the usual method of evaluation (mid-term and final examinations) interferes with the students' learning. Students tend to concentrate on memorizing information in order to pass the exam rather than exploring and experimenting with the new concepts and ideas that have been presented.

2. Assumptions about learning[1]. Let me offer a definition of learning. *Learning consists of evaluating new information in relation to information that is already understood and storing it in a form that is available for use in new situations.* The criterion that the information be available for use in *new* situations is an important one. Much of what is generally accepted as learning in our schools is actually memorization, and does not qualify as learning in the above sense. In fact, much of this memorization actually deters true learning.

Learning, as defined above, is a pleasurable experience. Human beings are inherently curious about their environment and enjoy learning about it. It is possible to provide classroom learning experiences that are both significant and enjoyable. And it is particularly important that future teachers be provided with enjoyable mathematical learning experiences if they are going to be able to make mathematics enjoyable for their own students. It is not sufficient, as I have often done, to entreat them fervently to do so. Most of them will not be successful unless we provide an appropriate model.

I make two assumptions about learning which indicate the conditions under which it occurs. *The first assumption is that students must be feeling good in order to learn.* When human beings are feeling bad their attention is directed toward past or present distress experiences and is not available for processing new information.

Now the teacher is not primarily a therapist and should not become preoccupied with that role. However, since it is useless and often harmful to attempt to present new information to a human who cannot evaluate it, it is worth the time and effort to develop and establish procedures that will make and keep the student's attention available for learning. In section 3, I describe the procedures I have adopted in this regard. For a more general discussion, see **[1]**.

Once a student's attention is available for learning, the second assumption gives us a guideline as to how to proceed.

The second assumption is that new information can be evaluated only when its relationship to information already assimilated is understood. To achieve this, new information must be presented in context and at the proper rate. This is the key to keeping the student's attention available for learning. This is because the effect on humans of having information presented out of context is usually much more extensive than their merely not being able to learn the new information. Such an experience is distressing in itself and, in addition, reminds us of past experiences of being defeated in attempts to learn.

3. The laboratory approach. The class met for one hour of lecture and three hours of laboratory each week. The lectures were devoted to providing perspective and historical commentary, answering questions, showing films and giving examinations. The laboratory period (there were two each week) provided the most significant learning experiences. In the laboratory the students were arranged into groups of four (preferably) or five. The groups were rearranged randomly at two to three week intervals.

Each period started with an activity designed to enable the students to know each other better and to help get their attention available for learning. For example, the students in the group would each, in turn, relate something they liked about themselves, or something good that had happened to them recently, or a memory of a pleasant mathematical learning experience. After a group had worked together for awhile, I had each person in turn tell the person on his/her right something they liked about that person. Although there was some resistance to these activities at first, the students eventually saw their value and enjoyed them.

The groups investigated the mathematical concepts using, as often as possible, manipulative equipment such as attribute blocks, Cuisenaire rods, geo-boards, tangrams, geo-blocks, and dice. Their investigation was directed by study guides designed to stimulate thought and discussion about

[1] This section is condensed from [1], which is also available from the author.

the various topics in the curriculum. The study guides varied somewhat depending on the topic under investigation but they usually consisted of questions, problems for investigation, explanations, and sometimes games. Often the group would be instructed to work in pairs with one person playing the role of the teacher who would explain a concept or an algorithm to the "student." The study guides were written with the small group approach in mind and so were designed to use effectively the resources of the small group situation. During the laboratory period the instructor and/or teaching assistant circulated from group to group, observing, providing additional stimulation when necessary and assisting the group with difficulties. Although reading was assigned from a textbook this will be eliminated in subsequent classes because sufficient written material will be added to the study guides.

It should be emphasized that although the laboratory equipment is designed for use by elementary school students, we used it in a much more sophisticated way. Even the most mathematically mature students found the course challenging. Both the mathematical concepts introduced and the expected level of understanding did not differ substantially from the previous courses I had taught using a lecture-textbook approach. It should also be noted that the attribute blocks, Cuisenaire rods, geo-boards and tangrams were available in versions suitable for individual use and purchase. The students were expected to purchase, share the purchase, or construct duplicates of these materials so that they could use most of the study guides at home as well as in the laboratory

Let us examine how the structure of the class relates to the assumptions about learning described in the previous section. The small group approach sets up a situation where the students are learning from each other. Such a situation reduces the fear of authority that interferes with learning. The communication of ideas is also of great benefit in the learning process. One learns best what one is teaching. Secondly, the introductory activities described above enable the students to get to know each other. This facilitates learning by reducing feelings of fear and alienation. In addition when the students relate positive experiences their attention is directed away from their distress and is thus more readily available for learning. Appreciation of the students by themselves and each other contradicts the insecurity that even the "best" students have about their abilities. Success at working together on problems and teaching mathematics to other students is also effective in contradicting the feelings of inadequacy. Furthermore, the small group approach gives the students more control over how they spend their time in class. This helps avoid the feelings of frustration or boredom that tend to arise when students are coerced into spending time studying ideas that don't interest them or are not allowed to spend time on the ideas that do interest them.

With regard to the second assumption, namely that new information must be presented in context, the small-group laboratory approach has some obvious advantages. The most obvious difficulty with the lecture method, both in theory and in practice, is the impossibility of presenting information in context for several people at one time. Each person in the classroom has had different mathematical experiences. Each person understands different concepts from previous lessons and courses. This suggests that one person teaching one other person is most effective. The small group approach approximates this and the reason for having four people in a group instead of two is to increase the resources available for each learning activity.

Another advantage of the small group approach is that the learner can assume greater responsibility for the rate of presentation of new information. Students are much more likely to ask questions and indicate areas of difficulty in small groups than in lecture situations.

There are also several advantages to using manipulative equipment in this course. The material is attractive and interesting in its own right. In addition, students immediately perceive its applicability for their own teaching. Even though we do not specifically investigate the procedures of teaching with the equipment, they are convinced that the mathematics they are learning is relevant to their future careers.

Another advantage to using the equipment is that it enables me to supply the students with mathematical experiences that they may have missed and that are necessary for their understanding more abstract mathematical concepts. One frequent difficulty in teaching mathematics is that the

students have not assimilated the necessary information to provide the context for the new information that we wish to teach. Very basic experiences with the physical world that are a prerequisite for understanding abstract mathematical concepts are often lacking and require special attention. The result of the absence of these experiences is that the student relies on memorization of techniques which will satisfy the teacher's demand for correct answers. This memorization is often not real learning in that the information may not be available for use in responding to new situations. Supplying some basic experiences with the mathematical-physical world to prospective teachers does more than facilitate their own learning, however. It will make it easier for them to encourage exploration and understanding instead of memorization of techniques and symbol manipulation.

My goal in evaluation is to maximize self-evaluation of learning and minimize evaluation by others unless specifically requested by the learner. People have the ability to assess their own learning. Evaluation by others undermines this ability and results in students being afraid to trust their own judgment. In addition, the resulting anxiety about being evaluated interferes significantly with learning. Since the desire to learn is inherent in every human, devices such as grades, tests, and awards should be unnecessary for motivation.

I have not yet achieved the goal of student self-evaluation. Complete realization of this goal would necessitate a considerable change in institutional procedure. I do plan on preparing a system of evaluation which would enable students to repeat attempts to demonstrate comprehension of learning units without penalty for previous failure.

4. Evaluation of the project. The two-quarter course, Mathematics for Elementary Teaching described above, and the same course using a lecture approach were pre-tested and post-tested using the Educational Testing Services Mathematics Basics Concepts Examination (STEP Series II). This multiple choice examination is designed to measure elementary mathematical concepts, abilities and skills. It was chosen because it is a standard examination based on the usual curriculum and would not give an advantage to the students who had participated in the laboratory class.

Both classes were given Form 1A during the first week of the first quarter of the class and both were given Form 1B during the last week of the second quarter of the class. We use throughout this description the converted scores determined by Educational Testing Service. This conversion corrects for differences in item difficulties between the forms by putting all scores on an equivalent scale.

The students who took both examinations will be referred to as the experimental group (for the laboratory class) and the control group (for the lecture class). The mean converted score of the experimental group (49 students) increased from 459.12 to 464.92, a difference of 5.80. The mean converted score of the control group (53 students) increased from 460.77 to 464.28, a difference of 3.51. The gain for the experimental group was greater; however, the difference was not statistically significant. (The mean gains in both groups represented an increase from approximately the 53rd percentile to the 63rd percentile.) The often expressed concern that the time devoted to non-academic activities, as described above, would have a deleterious effect on learning seems to be unfounded.

Now let us consider the attrition in these classes. The attrition is difficult to determine, or even define, since the course is a two-quarter course and some students do not complete the course in sequence. The one measure of attrition we do have is the percentage of students who having taken the first test (during the first week of the first quarter) did not take the second test (during the last week of the second quarter). The attrition in the laboratory class was 36% while the attrition in the lecture class was 51%. Although the measure used may not be absolutely accurate, the figures indicate that the attrition in the lecture class is substantially greater.

It is also interesting to investigate the pre-test scores of those students who did not take the post-test. We shall refer to these students as the laboratory attrition group and the lecture attrition group. The chart below compares the mean and median scores of these groups with the groups that did continue with the class.

CONVERTED SCORES ON PRE-TEST

	Mean	Median
Laboratory attrition group	456	453
Laboratory class	460	459
Lecture class	461	462
Lecture attrition group	463	464

It is seen that the mean and median scores of the laboratory attrition group on the pre-test is less, respectively, than the mean and median scores of the laboratory class, while for the lecture method the opposite is true.

This might lead one to the conjecture that if the higher-scoring students had remained in the lecture class, the gain for the lecture class would have been greater.

To examine this conjecture each class was divided into thirds based on the pre-test scores and the gains for each third were determined. The results are summarized below:

		Number of students	Mean score pre-test	Mean score post-test	Change
Students scoring in upper 1/3 on pretest	Laboratory class	16	473.88	475.19	+ 1.31
	Lecture class	18	470.78	468.94	− 1.84
Students scoring in middle 1/3 on pre-test	Laboratory class	15	459.27	468.27	+ 9.00
	Lecture class	18	461.06	465.94	+ 4.88
Students scoring in lower 1/3 on pre-test	Laboratory class	18	445.89	453.00	+ 7.11
	Lecture class	17	449.88	457.59	+ 7.71

The mean of the students scoring in the upper third on the pre-test actually decreased slightly in the lecture class, while increasing slightly in the laboratory class. So it is not likely that the attrition of the higher scoring students adversely affected the mean of the lecture class.

In conclusion these data, combined with the previous statements concerning the attrition rate and the nature of the students who do not continue, suggest that the small-group laboratory approach may be more successful than the lecture method in motivating those students with more mathematical knowledge and skills without sacrificing the education of the less prepared student.

Of course further study will be necessary to substantiate these claims and also to investigate the very important effect of the small-group laboratory approach on the attitude of the students toward mathematics. It seems clear however that the small-group laboratory approach used in this project does at least as well in improving mathematical performance on a standard mathematics examination. Any of the other benefits which are often claimed for the laboratory method (such as the promotion of problem-solving abilities, the encouragement of independent thinking, the development of understanding as opposed to memorizing, the development of ability to question and make conjectures) would be "icing on the cake." That some of these benefits are realized is suggested by the difference in attrition rates and by the fact that those who do drop from the lecture class tend to be among the more able students.

Reference

1. J. Weissglass, Feelings and mathematics: Responding to distress in the classroom, J. Math. Soc. Japan, 57(1975) Supplementary issue, Reports of Mathematical Education.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, SANTA BARBARA, CA 93106.

A MATHEMATICAL MODELING APPROACH TO INTRODUCTORY MATHEMATICS

GERARD P. PROTOMASTRO AND CECIL R. HALLUM

In this paper we shall discuss our experience with a new introductory mathematics course in our common curriculum at Loyola University. The common curriculum is Loyola University's general education requirement. Its purpose is to provide a liberalizing and humanizing dimension to the educational program of every undergraduate student. It is divided into lower- and upper-division courses with mathematics being considered as a lower-division course. In considering the inclusion of a mathematics requirement in the curriculum, it was decided that the course should not be a remedial one and that it should support the goals of the common curriculum. In the passage of this requirement, it was agreed upon that the course should be presented through a mathematical modeling approach — one which presents mathematics which is interesting, useful, and accessible. A student could test out of this requirement by receiving a percentile ranking of 80% or above on his SAT or ACT scores. Thus the course is taken mostly by students who have average or below average mathematics ability.

The goal of the course is to provide the student with an appreciation for and an understanding of the use of mathematics in other fields. The role of mathematical models in explaining and predicting phenomena in the real world is the central theme. The suggested syllabus, which can be adapted to meet a wide variety of freshman student profiles, is as follows:

MATHEMATICAL MODELING AND APPLICATIONS

I. Introduction and Philosophy of Model Building
A. Motivations for Model Building and Analysis (Aesthetic or Utilitarian).
B. Introductory Logistics and Its Utility.
II. Model Types
A. Purely Deterministic.
B. Static, or Deterministic with Simple Random Components.
C. Stochastic.
III. Model Construction and Analysis
A. Basic Principles.
B. Classic Examples (Modeling the Planetary System, Games of Chance, etc.).
C. Linear Optimization Models.
D. Modeling Growth Processes (Population growth, radioactive decay, epidemic modeling, etc.).
E. Stochastic Modeling (Markoff chains — applications to Behavioral Sciences).
IV. Practical Aspects of Model Building and Its Analysis
A. Simulation and the Role of the Computer.
B. Testing Basic Principles.

In addition to the above the student is responsible for the outside reading of articles on various aspects of mathematics. The essays assigned deal with mathematics both as a basic form of creative thought and as a tool employed in science and technology, in education, and indeed, throughout our social institutions. The purpose of these assigned readings is to make the student more culturally aware of mathematics by showing him what it is about mathematics that accounts for its expanding use and for the fascination which it holds for those who are proficient in it.

At the end of the course the students were asked to fill out a course evaluation form. It consisted of ten statements and the student had to indicate his agreement with the statement by using the following scale:

1 Strongly disagree, **2** Disagree, **3** Undecided, **4** Agree, **5** Agree strongly.

This scale was chosen because it is identical to the one used on the faculty evaluation forms distributed by the university to the students at the end of each semester. Our findings, based upon the result of the evaluation and upon our experience with the course, are contained in the following paragraph.

First and foremost the students definitely feel that the mathematical modeling approach illustrated to them the usefulness of mathematics in today's world, and that they now have a better

understanding of what mathematics is all about. They became aware of many aspects of mathematics with which they were unfamiliar and were impressed by the extent to which its forms of reasoning have spread into practically all aspects of modern life. Certainly this is an important consequence, since we are all aware of the continuing inadequacy of the public's understanding of how and to what extent the intricacies of mathematics affect contemporary human activities. Because mathematical models are best used as a motivational device for both reinforcing familiar concepts and introducing new subject matter, the students thought that the mathematics presented was very interesting. They strongly agreed that they preferred the study of mathematical modeling and its applications to a course mostly consisting of a repetition of high school topics presented in the usual way. The students believed that the outside readings supplemented the course by stressing the creative way mathematics comes to grips with nature and life in the modern world. Overall, the level of mathematics covered in this course is somewhat higher than that covered in a remedial one. But because of the motivational aspects of mathematical models, there was no significant difference in the final grades as compared to those of previous remedial courses.

In concluding this paper, we thoroughly recommend the mathematical modeling approach to introductory mathematics for the following reasons:

(1) Models enhance mathematics learning, making it more enjoyable and more interesting for both student and instructor.

(2) A mathematical model is a unifying concept for many diverse subjects.

(3) In general it is potentially very useful for mathematicians to comment on the interface between mathematics and the subjects to which it is applied.

(4) This approach makes mathematics more appealing to the students because they realize the ideas and tools of mathematics enable them to better understand the real world.

References

1. P. Abell, Model Building in Sociology, Schocken, New York, 1971.

2. N. Bailey, The Mathematical Theory of Epidemics, Hafner, New York, 1957.

3. M. Bartlett, Stochastic Population Models, Methuen, London, 1960.

4. R. Barton, A Primer on Simulation and Gaming, Prentice-Hall, Englewood Cliffs, N.J., 1970.

5. O. Bartos, Simple Models of Group Behavior, Columbia University Press, New York, 1967.

6. R. Bellman, K. Cooke and J. Lockett, Algorithms, Graphs and Computers, Academic Press, New York, 1970.

7. P. Benaurraf and H. Putnam, Philosophy of Mathematics, Selected Readings, Prentice-Hall, Englewood Cliffs, N.J., 1964.

8. I. Buchler and H. Notini, Game Theory in the Behavioral Sciences, University of Pittsburgh Press, Pittsburgh, 1969.

9. A. Charnes and W. Cooper, Management Models and Industrial Applications of Linear Programming, Wiley, New York, 1961.

10. G. Dantzig, Linear Programming and Extensions, Princeton University Press, Princeton, N.J., 1963.

11. W. Ewens, Population Genetics, Methuen, London, 1968.

12. W. Feller, An Introduction to Probability Theory and Its Applications, Wiley, New York, 1968.

13. H. Freudenthal, The Concept and the Role of the Model in Mathematics and Natural and Social Sciences, Gordon and Breach, New York, 1961.

14. D. Gale, The Theory of Linear Economic Models, McGraw-Hill, New York, 1960.

15. S. Gass, An Illustrated Guide to Linear Programming, McGraw-Hill, New York, 1970.

16. A. Glicksman, An Introduction to Linear Programming and the Theory of Games, Wiley, New York, 1963.

17. S. Goldberg, Probability, An Introduction, Prentice-Hall, Englewood Cliffs, N.J., 1960.

18. F. Harary, Graph Theory, Addison-Wesley, Reading, Mass., 1969.

19. F. Harary, R. Norman and D. Cartwright, Structural Models, Wiley, New York, 1965.

20. P. Hoel, Introduction to Mathematical Statistics, Wiley, New York, 1971.

21. J. Kemeny and J. Snell, Finite Markov Chains, Van Nostrand, Princeton, N.J., 1960.

22. J. Kemeny and J. Snell, Mathematical Models in the Social Sciences, M.I.T. Press, Cambridge, Mass., 1972.

23. J. Kemeny, H. Mirkil, J. Snell and G. Thompson, Finite Mathematical Structures, Prentice-Hall, Englewood Cliffs, N.J., 1959.

24. N. Keyfitz, Introduction to the Mathematics of Population, Addison-Wesley, Reading, Mass., 1968.
25. M. Kline, Mathematics in the Modern World, Freeman, San Francisco, Calif., 1968.
26. R. Kohn, A mathematical programming model for air pollution control, School Science and Mathematics, 69 (1969) 487–494.
27. P. Lancaster, Mathematics: Models of the Real World, Prentice-Hall, Englewood Cliffs, N.J., 1976.
28. R. Luce and H. Raiffa, Games and Decisions, Wiley, New York, 1957.
29. D. Maki and M. Thompson, Mathematical Models and Applications, Prentice-Hall, Englewood Cliffs, N.J., 1973.
30. J. Malkevitch and W. Meyer, Graphs, Models and Finite Mathematics, Prentice-Hall, Englewood Cliffs, N.J., 1974.
31. The Mathematical Sciences, M.I.T. Press, Cambridge, Mass., 1969.
32. D. Messick, Mathematical Thinking in Behavioral Sciences, Freeman, San Francisco, Calif., 1968.
33. A. Mizrahi and M. Sullivan, Mathematical Models and Applications: Suggestions for the High School Classroom, Math. Teacher, 66 (1973) 394–402.
34. J. Newman, The World of Mathematics, Simon and Schuster, New York, 1956.
35. I. Niven, Mathematics of Choice, Random House, New York, 1965.
36. O. Ore, Graphs and Their Uses, Random House, New York, 1963.
37. G. Owen, Game Theory, Saunders, Philadelphia, 1968.
38. F. Restle, Mathematical Models in Psychology, Penguin, Baltimore, 1971.
39. T. Saaty and F. Weyl, The Spirit and the Uses of the Mathematical Sciences, McGraw-Hill, New York, 1969.
40. R. Singleton and W. Tyndall, Games and Programs: Mathematics for Modeling, Freeman, San Francisco, Calif., 1974.
41. J. Smith, Mathematical Ideas in Biology, Cambridge University Press, London, 1968.
42. W. Stansfield, Genetics, Schaum's Outline Series, McGraw-Hill, New York, 1969.
43. G. Wimbish, Readings for Mathematics: A Humanistic Approach, Wadsworth, Belmont, Calif., 1972.

Department of Mathematics, Saint Peter's College, Jersey City, NJ 07306.
Department of Mathematical Sciences, Loyola University, New Orleans, LA 70118.

MISCELLANEA*

1. "It is remarkable that one can use integral calculus to obtain an intrinsic property of the prime numbers; but all mathematical truths are linked to each other, and all methods of discovering them are equally permissible."

—A. M. Legendre, Essai sur la théorie des nombres, 2nd ed., Paris, 1808, p. 398.

2. "On the other hand, researches based on the theory of integral equations are liable to give rise to uneasy feelings of suspicion in the mind of the ultra-orthodox mathematician."

—G. N. Watson, A Treatise on the theory of Bessel functions, Cambridge University Press, 1922, p. 578.

*** This department will publish, from time to time, interesting comments about mathematics, intriguing facts, and similar material. Contributions will be welcome and should be sent to the editor with appropriate references.**

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before August 31, 1977.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

E 2653*. *Proposed by Albert A. Mullin, Fort Hood, Texas*

A lattice point $(x, y) \in \mathbb{Z}^2$ is *visible* if $\mathrm{GCD}(x, y) = 1$. Prove or disprove: Given a positive integer n, there exists a lattice point (a, b) whose distance from every visible point is $\geqq n$.

E 2654. *Proposed by D. E. Daykin, Reading University, England*

Let $A = \{0, 1, 2, \ldots, n-1\}$. For $m \in A$ let $f(m, n)$ be the least integer k with the following property. If F is a family of subsets of A such that every $i \in A$ belongs to more than k members of F, then A can be covered by $n - m$ members of F. Evaluate $f(m, n)$ for $2m \leqq n$.

E 2655. *Proposed by Michael W. Chamberlain, University of Santa Clara*

Prove that for integral $n \geqq 2$ and $0 < x < n/(n+1)$ one has

$$(1 - 2x^n + x^{n+1})^n < (1 - x^n)^{n+1}.$$

E 2656. *Proposed by G. Tsintsifas, Thessaloniki, Greece*

Let $a_2, a_3, \ldots, a_n$ be positive real numbers and $s = a_2 + a_3 + \cdots + a_n$. Show that

$$\sum_{k=2}^{n} a_k^{1-1/k} < s + 2\sqrt{s}.$$

E 2657.* *Proposed by G. Tsintsifas, Thessaloniki, Greece*

Let $\mathcal{A} = A_0A_1 \ldots A_n$ and $\mathcal{B} = B_0B_1 \ldots B_n$ be regular simplices in $\mathbb{R}^n$. Assume that the ith vertex of $\mathcal{B}$ lies on the ith face of $\mathcal{A}$ $(0 \leqq i \leqq n)$. What is the minimal value of their similarity ratio λ ($\lambda\mathcal{A}$ congruent to $\mathcal{B}$; $\lambda > 0$)?

E 2658. *Proposed by W. Weston Meyer, General Motors Research Laboratories, Warren, Michigan*

(1) For $0<\alpha<\pi/2$ and integral $n\geqq 0$ show that

$$\int_0^\alpha \left(\frac{\sin\theta}{\sin\alpha}\right)^{2n} d\theta = \sum_{k=0}^{n} c_{nk} \int_0^\alpha \left(\frac{\tan\theta}{\tan\alpha}\right)^{2k} d\theta,$$

where the constants c_{nk} are independent of α.

(2) Find all polynomials P such that the ratio

$$\int_0^\alpha P\left(\frac{\sin\theta}{\sin\alpha}\right) d\theta \Big/ \int_0^\alpha P\left(\frac{\tan\theta}{\tan\alpha}\right) d\theta$$

is independent of $\alpha\in(0,\pi/2)$.

SOLUTIONS OF ELEMENTARY PROBLEMS

Diagonals in a 0–1 Matrix

E 2372 [1972, 773; 1973, 945]. *Proposed by E. T. H. Wang, University of Waterloo, Ontario, Canada*

Let A be an $n\times n$ matrix with entries zero and one, such that each row and each column contains precisely k ones. A generalized diagonal of A is a set of n elements of A such that no two elements appear in the same row or the same column. Show that A has at least k pairwise disjoint generalized diagonals, each of which consists entirely of ones.

II. *Recall of a solution.* Daniel B. Shapiro and S. David Shapiro observe that the published solution [1973, 945] is incorrect.

III. *Comments.* K. R. Rebman observes that the problem is well known and is equivalent to Theorem 5.3 in H. Ryser, *Combinatorial Mathematics*, p. 57 (Carus Math. Monograph 14).

Lowell W. Beineke and Raymond E. Pippert notice that the problem is equivalent to a well-known theorem in graph theory: Every finite regular bipartite graph is decomposable into 1-factors. For this classical result see D. König, *Theorie der Graphen*, Chelsea, 1950, p. 171.

A Special Group Operation on Natural Numbers

E 2574 [1976, 54]. *Proposed by F. David Hammer, Stockton State College, New Jersey*

Let $N=\{0,1,2,\ldots\}$ and let p be a prime. There is a binary operation $*$ on N satisfying $x*y\leqq x+y$ for all x,y in N such that $(N,*)$ is an abelian group with every element (except 0) of order p: for example, write x and y to base p and add individual digits mod p. Prove, or disprove, that this gives the only such operation.

Solution by Dmitri Nakassis, Washington, D.C. Under the hypothesis, $*$ is unique. Indeed, let $\circ$ be the operation mentioned in the proposal (addition of individual digits mod p), and let A_i be the set of all elements of N which are smaller than p^i. Then for every i, $i=0,1,2,\ldots$, $(A_i,\circ)$ is a group and $(A_i,*)=(A_i,\circ)$. That is, $(A_i,*)$ is a subgroup of $(N,*)$ and $*$ and $\circ$ coincide on A_i. This is trivially true if $i=0$.

Suppose then that it is also true for $i=k$. Let m stand for p^k and $m^{(j)}$ for the jth power of m in $(N,*)$. Since p is prime and m is not in A_k, the union

$$A_k\cup m*A_k\cup\cdots\cup m^{(p-1)}*A_k=S$$

is disjoint and each of the cosets $m^{(j)}*A_k$ has p^k elements. But we also have

$$m^{(j)}*n\leqq j\cdot p^k+n<p^{k+1} \tag{1}$$

for $0 \leqq j \leqq p-1$ and $n \in A_k$. Therefore $S = A_{k+1}$ and $(A_{k+1}, *)$ is a subgroup of $(N, *)$. Moreover, it follows from the inequalities (1) that

$$m^{(j)} * n = jm + n \tag{2}$$

for $0 \leqq j \leqq p-1$, $n \in A_k$. (This can be proved by double induction on n and j.)

Since both $(A_{k+1}, *)$ and $(A_{k+1}, \circ)$ are elementary abelian p-groups and $(A_k, *) = (A_k, \circ)$ it follows from (2) that also $(A_{k+1}, *) = (A_{k+1}, \circ)$.

Since the sequence $\{A_i\}$ is increasing, $\cup A_i = N$ and $(A_i, *) = (A_i, \circ)$ for $i \geqq 0$, it follows that $(N, *) = (N, \circ)$.

Also solved by David Bienenfeld (Israel), Steven Galovich, Donald Knuth, Donald Palmer, and Christiane Rollier & J. C. Binz (Switzerland).

A Non-symmetric Function

E 2575 [1976, 133]. *Proposed by David Shelupsky, City College of New York*

Solve the functional equation

$$f\left(\frac{x-y}{\log x - \log y}\right) = \frac{1}{2}f(x) + \frac{1}{2}f(y),$$

this to hold for all distinct x, y in $(0, \infty)$ and $f: (0, \infty) \to R$ to be continuous.

Solution by Albert Nijenhuis, University of Pennsylvania. Let f be a continuous solution of the given functional equation and let $g(x, y) = (x-y)/(\log x - \log y)$ for $x \neq y$ and $g(x, x) = x$.

Assume first that f is strictly monotonic. Then we have

$$f(x) + f(y) + f(u) + f(v) = f(g(x, y)) + f(g(u, v)) = f(g(g(x, y), g(u, v)))$$

and consequently the function

$$G(x, y, u, v) = g(g(x, y), g(u, v))$$

must be symmetric in x, y, u, v. This is not the case, because a computation shows that

$$G(2, 3, 5, 7) \doteq 3.95345724, \qquad G(2, 5, 3, 7) \doteq 3.95345474.$$

Hence f cannot be strictly monotonic. Therefore f has a local extremum at some point $a > 0$. The functional equation implies that $f(x) = f(a)$ in some neighborhood of a. Let A be the set consisting of all $b > 0$ such that f is constant on $[a, b]$. Then A is connected and hence is an interval. The functional equation implies that A is open and the continuity of f implies that A is closed in $(0, \infty)$, therefore $A = (0, \infty)$ and f is a constant.

Also solved by Irl Bivens, D. M. Bloom, Dana Kamerud, and R. Odoni (England).

Area of a Projection of an Ellipsoid

E 2576 [1976, 133]. *Proposed by Robert L. Helmbold, Northridge, California*

What is the area $A(\mathbf{n})$ of the orthogonal projection of the ellipsoid $(x/a)^2 + (y/b)^2 + (z/c)^2 = 1$ onto a plane perpendicular to the unit vector $\mathbf{n} = (n_1, n_2, n_3)$?

Solution by I. I. Kolodner, Carnegie-Mellon University. Let V be the volume of the cylinder slice C which has for base the said projection on the plane $P: \mathbf{n} \cdot \mathbf{x} = 0$ and $\mathbf{n}$ for generator. Then $A(\mathbf{n}) = V$. Subjecting the space to the affine transformation $\mathbf{x} \to \mathbf{x}' = T\mathbf{x}$ with $T = \operatorname{diag}(a^{-1}, b^{-1}, c^{-1})$, the ellipsoid maps into the unit ball $|\mathbf{x}'| = 1$, the cylinder C maps into cylinder C' with volume

$V' = V \det T = V/abc$, while $\mathbf{n}$ maps to $\mathbf{n}' = T\mathbf{n}$. We now observe that $V' = \pi|\mathbf{n}'|$, so it follows that

$$A(\mathbf{n}) = V = abcV' = \pi abc\,[(n_1/a)^2 + (n_2/b)^2 + (n_3/c)^2]^{1/2}.$$

Also solved by Leon Gerber, Clark Givens, Michael Goldberg, Charles Grosch, H. Kappus (Switzerland), Jack Kuipers, O. P. Lossers (Netherlands), John Maddocks (Scotland), E. Trost (Switzerland), University of South Alabama Problem Group, and the proposer.

Editor's Comments. Gerber proves the analogous result for the n-dimensional ellipsoid $\Sigma_{i=1}^{n}(x_i/a_i)^2 = 1$. If $\mathbf{p} = (p_i)$ is the unit vector then the corresponding $(n-1)$-dimensional area is

$$A(\mathbf{p}) = c_{n-1} d\, a_1 a_2 \cdots a_n,$$

where $d^2 = \Sigma(p_i/a_i)^2$ and

$$c_n = \pi^{n/2}/\tfrac{1}{2}n\Gamma(\tfrac{1}{2}n).$$

Grosch notes that if E is the ellipse obtained by intersecting the given ellipsoid by the plane $n_1x + n_2y + n_3z = 0$ then its area is equal to

$$\pi abc(n_1^2a^2 + n_2^2b^2 + n_3^2c^2)^{-1/2}.$$

Several contributors computed the area of E instead of the required area.

Restricted Ménage Numbers

E 2577 [1976, 133]. *Proposed by F. W. Light, Jr., Jonestown, Pennsylvania*

Given the $2 \times n$ Latin rectangle

$$\begin{matrix} 1 & 2 & 3 & \dots & n-1 & n \\ 2 & 3 & 4 & \dots & n & 1 \end{matrix}$$

find the number of ways $f(k)$ in which a $3 \times n$ Latin rectangle can be built up from it by adding a third row starting with k, where k is one of the numbers $3, 4, \dots, n$. (If the initial number a is not prescribed, the number of possible third rows is, of course, the already known ménage number U_n.)

Solution by the proposer (revised by the editor). Let σ_i^j be the operator which acts on $(0,1)$-matrices by replacing the (i,j)-th entry, say x, by $1-x$. Let θ_i^j be the operator which acts on $(0,1)$-matrices by deleting the ith row and the jth column. If A and B are $(0,1)$-matrices we denote by $A \vee B$ the matrix

$$\left(\begin{array}{c|c} A & J_1 \\ \hline J_2 & B \end{array}\right)$$

where J_1 and J_2 have all entries equal to 1.

If A is a $(0,1)$-matrix and its (i,j)-th entry is 1 then for permanents we have

$$\text{(1)} \qquad \operatorname{per}(A) = \operatorname{per}(\sigma_i^j A) + \operatorname{per}(\theta_i^j A).$$

Let C_n $(n \geqq 2)$ be the cyclic $n \times n$ matrix whose first row is $0, 1, 1, \dots, 1, 1, 0$. The ménage numbers are given by $U_n = \operatorname{per}(C_n)$, $n \geqq 2$. We let $A_n = \sigma_1^n C_n$ and $x_n = \operatorname{per}(A_n)$, $n \geqq 2$.

Since $\theta_1^n A_n = A'_{n-1}$ (prime is the transposition operator), (1) gives $\operatorname{per}(A_n) = \operatorname{per}(C_n) + \operatorname{per}(A'_{n-1})$, i.e., $x_n = U_n + x_{n-1}$ $(n \geqq 3)$. Therefore

$$\text{(2)} \qquad x_m = \sum_{i=3}^{m} U_i \qquad (m \geqq 2).$$

The required number $f(k)$ or, more precisely, $f(n,k)$ is equal to $\operatorname{per}(\theta_k^1 C_n)$. By permuting

cyclically the rows of $\theta_k^1 C_n$ so that its first row becomes the last, we obtain the matrix

$$\theta_{k-1}^1 C_n' = A_{k-2} \vee A_{n-k+1}', (A_1 = (0)).$$

Hence

$$f(n, k) = \operatorname{per}(A_{k-2} \vee A_{n-k+1}) = \operatorname{per}(\sigma_{k-1}^{k-2} A_{n-1}). \tag{3}$$

It is clear from (3) that

$$f(n, k) = f(n, n - k + 3), \tag{4}$$

and thus we can assume that $2k \leqq n + 3$.

Since

$$\theta_{k-1}^{k-2} A_{n-1} = \sigma_{k-2}^{k-2} A_{n-2}, \quad \theta_{k-2}^{k-2} A_{n-2} = \sigma_{k-2}^{k-3} A_{n-3}$$

we obtain from (3) and (1) that

$$\begin{aligned} f(n, k) &= x_{n-1} + \operatorname{per}(\theta_{k-1}^{k-2} A_{n-1}) \\ &= x_{n-1} + x_{n-2} + \operatorname{per}(\sigma_{k-2}^{k-3} A_{n-3}). \end{aligned} \tag{5}$$

From (5) and (3) it follows that

$$f(n, k) = \sum_{i=n+3-2k}^{n-1} x_i \qquad (2k \leqq n + 3). \tag{6}$$

Of course, (2) and (6) can be used to express $f(n, k)$ in terms of U_i.

Also solved by Anthony Barkauskas, J. C. Binz (Switzerland), O. P. Lossers (Netherlands), and Reinhard Razen (Austria).

Polynomials Reducible Modulo Every Prime

E 2578 [1976, 133]. *Proposed by Carl Pomerance, University of Georgia*

Prove that $x^4 + 1$ is reducible over every field of prime characteristic. Do the same for $x^4 - x^2 + 1$.

Solution by Lindsay Childs, State University of New York, Albany. The following theorem gives a more general result.

THEOREM. *Given any prime p and integers a, b, the polynomial $P(x) = x^4 + ax^2 + b^2$ is reducible mod p.*

This theorem is essentially Lemma 12 of B. Fein and M. Schacher, *Solutions of pure equations in rational division algebras* I, J. Algebra 17 (1971), 83–93. They attribute the result and the following proof to Hilbert.

Proof. If $p = 2$ the claim is true because $P(x)$ is one of the following polynomials: x^4, $x^2(x^2 + 1)$, $(x + 1)^4$, $(x^2 + x + 1)^2$.

Suppose p is odd and let s be such that $a \equiv 2s \pmod{p}$. Then

$$\begin{aligned} P(x) &\equiv x^4 + 2sx^2 + b^2 \equiv (x^2 + s)^2 - (s^2 - b^2) \qquad \pmod{p} \\ &\equiv (x^2 + b)^2 - (2b - 2s)x^2 \equiv (x^2 - b)^2 - (-2b - 2s)x^2. \end{aligned}$$

It suffices to show that one of $s^2 - b^2$, $2b - 2s$, $-2b - 2s$ is a square mod p. If neither $2b - 2s$ nor $-2b - 2s$ is a square then their product $4(s^2 - b^2)$ is a square mod p. Hence $s^2 - b^2$ is also a square mod p.

Also solved by James Alonso, Anders Bager (Denmark), Madelaine Bates, Marc Berger, L. C. Bourburgh, David Buchthal, Peter de Buda, A. Charnow, H. M. Edgar, Lorraine Foster, Ira Gessel, Richard Gibbs, Robert Gilmer, Solomon Golomb, M. G. Greening (Australia), Frederick Hoffman, Herbert Holden, Frederick Humburg, A. A. Jagers (Netherlands), Erwin Just, Victor Keiser, Kwangil Koh & Tiang Luh, L. Kuipers (Switzerland), S. C. Locke (Canada), O. P. Lossers (Netherlands), Joseph Muskat (Israel), Keith Nicholson (Canada), Hugh Noland, Michael Nutt (Canada), Bob Prielipp, James Ridley (South Africa), Eric Rosenthal, Nan-Shan Shou (Hong Kong), Paul Smith (Canada), Gregg Testini (Thailand), Stephen Tillman, Ernst Trost (Switzerland), William Vélez, Roger Weitzenkamp, E. T. Wong, and the proposer.

Editor's Comments. Almost all solvers notice that the given polynomials are the cyclotomic polynomials $F_8(x)$ and $F_{12}(x)$. It is easy to deduce from a theorem of W. J. Guerrier (this MONTHLY, 75 (1968), p. 46) that the cyclotomic polynomial $F_n(x)$ is reducible modulo every prime if and only if the group of units of the ring $Z/(n)$ is not cyclic. It is well known that this group is cyclic if and only if $n = 1, 2, 4, p^k$ or $2p^k$ where p is an odd prime. The reference to Guerrier was supplied by Gilmer. He also points out that the cyclotomic polynomials $F_n(x)$ for $n = 39, 55, 95, 111$ are reducible modulo every integer $m \geqq 2$. For this he refers to his paper in Canad. Math. Bull. 16 (1973), 521–523. Bager remarks that the second part of this problem was posed by him in Elemente der Math. 12 (1957), p. 46 with a solution in 13 (1958), pp. 87–88. Shou points out that both claims of the problem are particular cases of a theorem of R. Swan, Pacific J. Math. 12 (1962), 1099–1106.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Hill Center, Rutgers University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate signed sheets and should be mailed before August 31, 1977.

6150. *Proposed by Albert A. Mullin, Fort Hood, Texas*

Let G be any groupoid. Call $e \in G$ a *near identity* of G if e is idempotent and $ex = xe = x$ *fails* for at most one $x \in G$. It is well known that G can have at most one identity. (1) Show that if G is a semigroup then it can have at most two near-identity elements. Every group has precisely one near-identity element. (2) Give an example of an *uncountably* infinite semigroup with precisely two near identities which contains a *countably* infinite semigroup with precisely two near identities.

6151. *Proposed by Clarence H. Best, Central Missouri State University*

A two-dimensional array is defined according to the following rule:

$$a_{11} = 1,\ a_{i1} = a_{1,i-1}\,(i > 1),\ a_{ij} = a_{i+1,j-1} + a_{i,j-1}\,(j > 1).$$

(A) Prove that a_{1j} equals the number of distinct partitions of a j-element set.
(B) Choose an nth order determinant D_n from the upper left corner of the array and prove

$$D_n = \prod_{0 \leqq i \leqq n-1} i!.$$

6152. *Proposed by R. Raphael, Université de Poitiers, France*

In some rings one has unique factorization for ideals. Show that the following limited form of factorization holds in all rings: if I_j, $j = 1, \ldots, n$ are distinct non-zero ideals in a ring R, and if a_j and b_j are positive integers with $a_j < b_j$ for each j then

$$\prod_{j=1}^{n} I_j^{a_j} = \prod_{j=1}^{n} I_j^{b_j} \quad \text{implies} \quad \prod_{j=1}^{n} I_j^{a_j} = \prod_{j=1}^{n} I_j^{c_j},$$

where c_j, $j = 1, \ldots, n$, are any integers satisfying $a_j \leqq c_j \leqq b_j$. In particular $\Pi I_j^{a_j} = \Pi I_j^{a_j+1}$. Show by an example that this is best possible, that is, show that one can have the products equal when the exponents are not.

6153. *Proposed by Bernardo Mz.-Recamán, University of Warwick, England*

Let $\pi(x)$ denote the number of primes $\leqq x$. Are there infinitely many integers, such as 2, 4, 6, 8, 30, 33, 100, with the property that $\pi(n)$ divides n?

6154. *Proposed by Richard Stanley, Massachusetts Institute of Technology*

Define a sequence of polynomials (with rational coefficients) as follows: $p_0(x)=1$, $p_n(0)=0$ if $n>0$, and $p'_{n+1}=p_n(1-x)$ if $n>0$. Thus $p_1(x)=x$, $p_2(x)=x-\frac{1}{2}x^2$, $p_3(x)=\frac{1}{2}x-\frac{1}{6}x^3$, etc. Find $p_n(x)$. In particular, what is $p_n(1)$?

6155. *Proposed by Milton P. Eisner, Ball State University*

Let $\{x_1, x_2, \ldots, x_k\}$ be a set of numbers. Define the *width* of the set to be $\min_{i\neq j}\{|x_i - x_j|\}$. Suppose the k numbers are selected at random from the set $\{1, 2, \ldots, n\}$. Find the expected value of the width of the resulting set if the numbers are chosen without replacement.

SOLUTIONS OF ADVANCED PROBLEMS

$\Pi \neq 83^3$

6044 [1975, 766]. *Proposed by Jacques Gilles, Saint-Servais, Belgium*

Show that $\Pi(\alpha^4+\alpha+1)=83^3$, the product being taken over all the roots of $\alpha^{49}=1$ (except $\alpha=1$).

I. *Solution by Jon L. Sicks, University of Massachusetts.* For the moment, do not except $\alpha=1$, but rather consider the function $f(\alpha_1,\ldots,\alpha_n)=\Pi_{\lambda=1}^{n}(\alpha_\lambda^4+\alpha_\lambda+1)$, where the α_λ are the roots of $x^n-1=0$; viz. $\alpha_\lambda=\exp(2\pi\lambda i/n)$. f is a symmetric polynomial with integral coefficients in α_1, $\alpha_2,\ldots,\alpha_n$ and hence may be expressed as a polynomial with integral coefficients in the elementary symmetric functions $\sigma_1,\ldots,\sigma_n$ of $\alpha_1,\ldots,\alpha_n$. (See, e.g., p. 442, *Algebra*, by G. Chrystal, Chelsea, 1964.)

But the coefficient of the $(n-k)$th term of x^n-1 is $(-1)^k\sigma_k$, so the σ_k are all integers. Hence $f(\alpha_1,\ldots,\alpha_n)$ is an integer.

Computing $f(\alpha_1,\ldots,\alpha_{49})$, using a computer, we find

$$\text{re } f(\alpha_1,\ldots,\alpha_{49})=15117188.999973$$

$$\text{im } f(\alpha_1,\ldots,\alpha_{49})=-1.6760602695047\times 10^{-4}.$$

Knowing that $f(\alpha_1,\ldots,\alpha_{49})$ is an integer, we conclude that $f(\alpha_1,\ldots,\alpha_{49})=15117189=3\times 197\times 25{,}579$, so the quantity asked for in the problem is $1/3\, f(\alpha_1,\ldots,\alpha_{49})=197\times 25579$, not 83^3.

II. *Solution by Lorraine Foster, California State University, Northridge.* Let ε denote a primitive 7th root of unity over the rational field Q. Define the polynomial $n(x)=x^4+x+1$. Also define

$$P_1=\prod_{i=1}^{6} n(\varepsilon^i),\ P_2=\prod_{i=1}^{6}\prod_{\alpha^7=\varepsilon^i} n(\alpha).$$

Then $P=P_1P_2$. Suppose that θ is a root of $x^7-\varepsilon=0$. Then θ is one of the 42 primitive 49th roots of unity and $[Q(\theta)\colon Q]=42$ as that $[Q(\theta)\colon Q(\varepsilon)]=7$ and $x^7-\varepsilon$ is irreducible in $Q(\varepsilon)[x]$. Therefore

$$\prod_{\alpha^7=\varepsilon} n(\alpha)={}_{Q(\theta)}N_{Q(\varepsilon)}(n(\theta)),$$

(where in general ${}_EN_F(\tau)$ will denote the norm of τ in the extension field E of F). Now ${}_{Q(\theta)}N_{Q(\varepsilon)}(n(\theta))=\det A$ where $n(\theta)X=AX$, $X=[1,\theta,\ldots,\theta^6]^T$, and

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ \varepsilon & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & \varepsilon & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & \varepsilon & 0 & 0 & 1 & 1 \\ \varepsilon & 0 & 0 & \varepsilon & 0 & 0 & 1 \end{bmatrix}$$

Define $g(\varepsilon) = \det A = \varepsilon^4 - 7\varepsilon^3 + 14\varepsilon^2 - 6\varepsilon + 1$. Then

$$P_2 = \prod_{i=1}^{6} g(\varepsilon^i) = \prod_{i=1}^{3} g(\varepsilon^i) g(\varepsilon^{-i}).$$

Define $\gamma = g(\varepsilon)g(\varepsilon^{-1})$. Since $\varepsilon^7 = 1$, $\varepsilon^6 = -\Sigma_{i=0}^{5} \varepsilon^i$, we find

$$\gamma = 478 + 265(\varepsilon^2 + \varepsilon^{-2}) + 183(\varepsilon^3 + \varepsilon^{-3}) = 112 + 265\delta + 183\delta^2$$

where $\delta = \varepsilon^2 + \varepsilon^{-2}$. Thus $P_2 = {}_{Q(\delta)}N_Q(\gamma)$. By inspection we find the minimum function of δ over Q to be $m(x) = x^3 + x^2 - 2x - 1$. Therefore $\gamma Y = BY$ where $Y = [1, \delta, \delta^2]^T$,

$$B = \begin{bmatrix} 112 & 265 & 183 \\ 183 & 478 & 82 \\ 82 & 347 & 396 \end{bmatrix}$$

$P_2 = \det B = 5039063$.

To evaluate P_1, observe that $n(\varepsilon)n(\varepsilon^{-1}) = \delta(\delta - 1)$, so that

$$P_1 = {}_{Q(\delta)}N_Q(\delta(\delta-1)) = {}_{Q(\delta)}N_Q(\delta)\,{}_{Q(\delta)}N_Q(\delta - 1)$$
$$= (-m(0))(-m(1)) = 1.$$

Thus $P = 5039063$.

III. *Solution by James G. Mauldon, Amherst College.* One method of proving that the given product (P say) is not equal to 83^3 is to observe that the polynomial $P = \Pi_1^{48}(1 + x^i + x^{4i})$ is divisible by the polynomial $\Sigma_0^6 x^{7i}$ Then, taking $x = 1$, we have $P \equiv 3^{48} \equiv 1 \pmod 7$ whereas $83^3 \equiv -1 \pmod 7$.

Also solved by W. O. Alltop, Ben Bowen, Paul Bruckman, David Carlson, Bomshik Chang (Canada), James Cross & Howell Herring, Veikko Ennola & K. Inkeri (Finland), Douglas Flegal, Benedict Freedman, F. Göbel & J. A. van Hulzen (Netherlands), Charles Hampton, Harvard Mathematics Department Common Room, Mike Hendy (New Zealand), I. M. Isaacs, Allan Johnson, Roger Kingsley, D. H. Lehmer, R. M. Mathsen, J. G. Mauldon (two solutions), L. E. Mattics, M. J. Pelling (Nigeria), Otto Ruehr, Jonathan Ryshpan, William Watkins, John White, Philip Young, and Aleksandras Zujus.

Editor's Notes. (1) With $P_n = \Pi_{k=1}^{n-1}(1 + \omega^k + \omega^{4k})$, $\omega = \exp(2\pi i/n)$, $n = 2, 3, 4, \ldots$, Bruckman develops several methods for calculation, and from a table of values is led to the following conjectures:

(A) $2^4 | P_n$ iff $n \equiv 0 \pmod{15}$;

(B) $2^8 | P_n$ iff $n \equiv 0 \pmod{30}$;

(C) $3^3 | P_n$ iff $n \equiv 0 \pmod{13}$;

(D) $3^7 | P_n$ iff $n \equiv 0 \pmod{39}$;

(E) $5 | P_n$ iff $n \equiv 0 \pmod 4$;

(F) $5^2 | P_n$ iff $n \equiv 0 \pmod{20}$;

(G) $7 \nmid P_n$ for all n;

(H) $11 | P_n$ iff $n \equiv 0 \pmod{10}$.

(2) Ennola and Inkeri show that P_n is a cube only for $n = 13$ if $n \leqq 51$. Ruehr verifies this up to $n = 60$.

(3) Many computer solutions — complete programs and print-outs — were received. Mauldon's two solutions using APL can be used to support the contention that to date APL is the greatest!

(4) This problem brought forth many interesting comments from several solvers — e.g., "in the hands of a good carpenter, what is an inch, more or less?"

Analytic Functions

6045 [1975, 766]. *Proposed by J. B. Rosser, University of Wisconsin*

Let D be a domain of the complex plane. For each fixed a let D_a be the set of z's such that both $a+z$ and $a-z$ lie in D. Choose a fixed complex α and let $f(z)$ be a function such that for each fixed a

$$f(a+z)+\alpha f(a-z)$$

is analytic in D_a.

Can one conclude that $f(z)$ is analytic throughout D? If not, give some additional weak conditions on f from which one could infer this.

Solution by Gerd H. Fricke, Wright State University. Let $F_a(z) = f(a+z)+\alpha f(a-z)$ be analytic in D_a for any $a \in D$. If $\alpha^2 \neq 1$ then $f(z)$ is analytic on D. If $\alpha = -1$ and $f(z)$ is bounded on compact subsets of D then $f(z)$ is analytic on D. If $\alpha = 1$ and $f(z)$ is bounded on compact subsets of D then $f(z)$ is continuous but need not be analytic.

Proof. (a) Let $\alpha^2 \neq 1$. Then $F_a(z) - \alpha F_a(-z) = (1-\alpha^2)f(a+z)$ is analytic on D_a. Thus by varying a, and because $1-\alpha^2 \neq 0$, $f(z)$ is analytic on D.

(b) Let $\alpha = -1$ and let $f(z)$ be bounded on compact subsets of D. Let $F_a(z) = f(a+z) - f(a-z)$ and choose a in D and $r > 0$ such that $\{z : |z-a| \leqq 2r\} \subset D_a$. Let

$$M = \sup_{|z-a|\leqq 2r} \{|f(z)|\}.$$

Thus, for any $|b-a| < r$

$$\frac{|F_b^{(n)}(0)|}{n!} \leqq \max_{|z-b|=r} \{|F_b(z)|\} r^{-n} \leqq \frac{2M}{r^n} \quad \text{for} \quad n = 0,1,\ldots.$$

Hence, since $F_b(z)$ is odd,

$$|f(a+h)-f(a)| = |F_{a+\frac{1}{2}h}(\tfrac{1}{2}h)| \leqq \sum_{n=0}^{\infty} \frac{2M}{r^{2n+1}} \left|\frac{h}{2}\right|^{2n+1}$$

which converges to zero for h converging to zero. Thus $f(z)$ is a continuous function.

Let R be a rectangle in $|z-a| < r$ with center b and vertices w_1, w_2, w_3, and w_4. Let ∂R denote the boundary of R oriented counterclockwise. Let $z_i = w_i - b$ for $i = 1,2,3,4$. Then the z_i's are the vertices of a rectangle R_1 in $|z| < r$ with center at the origin. Let ∂R_1 denote the boundary of R_1 (oriented counterclockwise). Thus

$$\begin{aligned} 0 = \int_{\partial R_1} F_b(z)dz &= \int_{\partial R_1} f(b+z)dz - \int_{\partial R_1} f(b-z)dz \\ &= \int_{\partial R_1} f(b+z)dz + \int_{\partial R_1} f(b+w)dw \\ &= 2\int_{\partial R_1} f(b+z)dz = 2\int_{\partial R} f(z)dz. \end{aligned}$$

Hence $\int_\gamma f(z)dz = 0$ for any closed rectangular path in $\{z : |z-a| < r\}$ and therefore $f(z)$ is analytic on $|z-a| < r$ and thus analytic on D.

(c) Let $\alpha = 1$ and let $f(z)$ be bounded on compact subsets of D. Now

$$\begin{aligned}2[f(a+h)-f(a)] &= f(a+2h)-f(a)-f(a+2h)-f(a)+2f(a+h)\\ &= f(a+2h)-f(a)-[F_{a+h}(h)-F_{a+h}(0)].\end{aligned}$$

Thus $f(a+h)-f(a)$

$$= 2^{-k}[f(a+2^k h)-f(a)] - \sum_{n=1}^{k} 2^{-n}[F_{a+2^{n-1}h}(2^{n-1}h)-F_{a+2^{n-1}h}(0)].$$

By the same argument as in (b)

$$\frac{F_b^{(n)}(0)}{n!} \leqq \frac{2M}{r^n}.$$

Hence, for $|2^{2k}h| < r$,

$$\begin{aligned}|f(a+h)-f(a)| &\leqq 2^{-k}2M + \sum_{n=1}^{k} 2^{-n}\sum_{j=1}^{\infty}\frac{2M}{r^{2j}}(2^{n-1}h)^{2j}\\ &\leqq 2^{1-k}M + \sum_{n=1}^{k} 2^{-n}\sum_{j=1}^{\infty} 2M(2^{-k-1})^{2j}\\ &\leqq 2^{1-k}M + \sum_{n=1}^{k} 2^{-n}M2^{-k}\\ &\leqq 2^{1-k}M + 2^{-k}M \leqq 2^{2-k}M\end{aligned}$$

which converges to zero for k tending to infinity. Thus f is continuous.

To show that $f(z)$ need not be analytic, consider $f(z) = z^2 + \operatorname{Re} z$; $f(z)$ is continuous, not analytic and $F_a(z) = 2\operatorname{Re} a + (a-z)^2 + (a+z)^2$ is an analytic (non-constant) function on $\mathbb{C}$.

Also solved by Paul Chernoff and by the proposer.

Editor's Note. In the case $\alpha = 1$, Chernoff shows that $f(z)$ may be written $\varphi(z) + h(z)$ where φ is analytic and h is an additive function.

Comparing Decompositions of Polynomials

6046 [1975, 766]. *Proposed by Stephen McAdam, University of Texas, Austin*

Let f and g be two nonconstant monic irreducible polynomials over the field K. Let u and v be roots of f and g respectively in some extension field of K. Suppose that over $K[v]$, the irreducible decomposition of f is $f = f_1^{e_1}, \ldots, f_n^{e_n}$ while over $K[u]$, g decomposes into $g = g_1^{d_1}, \ldots, g_m^{d_m}$. Then $n = m$ and, when appropriately ordered, $e_i = d_i$ and $\deg g_i/\deg f_i = \deg g/\deg f$.

Solution by Robert Gilmer, Florida State University. Let X and Y be indeterminates over K and let A be the ideal of $K[X, Y]$ generated by $f(X)$ and $g(Y)$. The ring isomorphisms taken in the following sequence are the canonical isomorphisms.

$$\begin{aligned}K[X,Y]/A &\simeq [K[X,Y]/(f(X))]/[A/(f(X))]\\ &\simeq K(u)[Y]/(g(Y)) \simeq K(u)[Y]/(g_1^{d_1}\cdots g_m^{d_m})\\ &\simeq K(u)[Y]/(g_1^{d_1}) \oplus \cdots \oplus K(u)[Y]/(g_m^{d_m}).\end{aligned}$$

Similarly we obtain the isomorphism

$$K[X, Y]/A \simeq K(v)[X]/(f_1^{e_1}) \oplus \cdots \oplus K(v)[X]/(f_n^{e_n}).$$

For each i between 1 and m, the ring $R_i = K(u)[Y]/(g_i^{d_i})$ is a principal ideal ring with unique maximum ideal $(g_i)/(g_i^{d_i})$, which is nilpotent of order d_i; moreover, the residue field of R_i is, to within isomorphism, $K(u)[Y]/(g_i)$ — an extension field of K of degree $(\deg f)(\deg g_i)$ over K. Corresponding statements apply to the rings S_j, where $S_j = K(v)[X]/(f_j^{e_j})$ for each j between 1 and n. Therefore

$$R_1 \oplus \cdots \oplus R_m = S_1 \oplus \cdots \oplus S_n$$

under an isomorphism that induces the identity mapping on K. Since the rings R_i and S_j are nonzero and indecomposable, it is then known (Theorem 3, page 205, of Zariski and Samuel, *Commutative Algebra*, Volume I) that $m = n$ and that by proper ordering the rings R_i and S_j are K-isomorphic for each i between 1 and n. From this isomorphism it follows that $d_i = e_i$ and $(\deg f)(\deg g_i) = (\deg g)(\deg f_i)$ for each i. This completes the solution of the problem.

Also solved by D. Alamelu (India), Douglas Costa, and the proposer.

Conformal Maps of Ellipses onto Ellipses

6047 [1975, 766]. *Proposed by C. D. Minda, University of Cincinnati*

Let E_1 and E_2 be ellipses in the complex plane. Prove that there is a conformal mapping of the interior of E_1 onto the interior of E_2 which maps the foci of E_1 onto the foci of E_2 if and only if E_1 and E_2 have the same eccentricity. Moreover, show that if such a conformal mapping exists, then it must necessarily be of the form $az + b$ for some complex numbers a and b with $a \neq 0$.

Solution by Paul R. Chernoff, University of California, Berkeley. Consider the ellipse $E(e)$ with eccentricity e and foci at ± 1; its equation is $e^2x^2 + e^2y^2/(1-e^2) = 1$. The conformal map

$$w = \left(\frac{1+\sqrt{1-e^2}}{2e}\right)z + \left(\frac{1-\sqrt{1-e^2}}{2e}\right)z^{-1}$$

maps the annulus $A(e)$: $(1-\sqrt{1-e^2})/e < |z| < 1$ onto the region between the ellipse $E(e)$ and the slit $-1 \leqq x \leqq 1$.

Now suppose that f is a linear conformal map between two ellipses taking the foci of one to the foci of the other. By making linear transformations we may assume that the two ellipses are $E(e_1)$ and $E(e_2)$ and that f fixes the foci ± 1. Since f is uniquely determined, reflection shows that f is real on the real axis, and therefore f maps the segment from -1 to 1 onto itself.

Then clearly, via f, we get a conformal mapping from the annulus $A(e_1)$ onto the annulus $A(e_2)$. But it is well known that two annuli are conformally equivalent if and only if they have the same ratios of inner to outer radii. It follows that $e_1 = e_2$.

Also solved by M. S. Klamkin (Canada), G. R. Padmanabhan (India), David Styer, and the proposer.

Harmonic Numbers

6048* [1975, 856]. *Proposed by H. M. Edgar, San Jose State University*

A positive integer n is said to be harmonic (see O. Ore, *On the averages of the divisors of a number*, this MONTHLY, December 1948, pp. 615–619) if the ratio $n\tau(n)/\sigma(n)$ is again integral. (Here $\tau(n)$ denotes the number of positive integral divisors of n and $\sigma(n)$ denotes their sum.)

(a) Are there any harmonic numbers other than the number one which are perfect squares?
(b) Do there exist infinitely many harmonic numbers?

Comment by Thomas E. Elsner, General Motors Institute. (a) An answer is not available; but by search no such square is less than 10^6. Some partial results are established: The residue for the quotient on any power $mp-r$ of prime p with $0 \leqq r < p$ is

$$m + (1-r)p^{mp-r}$$

and so no power of a prime is harmonic and the best possible residue is 1. There are an infinite number of "nearly harmonic" squares since for all odd primes the square p^{p-1} has residue 1. Another question would concern the existence of any other n with harmonic residue 1.

(b) Every perfect number is harmonic, but this proves to be of little help. In fact the harmonic numbers seem to be very closely linked to perfect numbers, with no odd harmonic other than 1 showing itself. Between the nth and $(n+1)$st perfect number there are 2^{n-1} other harmonic numbers for $n = 1, 2, 3$. If this pattern continues, one problem appears to be as difficult as the other.

Subgroups of the Symmetric Group

6049 [1975, 856]. *Proposed by D. E. Knuth, Stanford University*

What group is generated by the two cyclic permutations $(1, 2, \ldots, m)$ and $(1, 2, \ldots, n)$ when $1 < m < n$?

Solution by J. C. Lagarias, Bell Laboratories, Murray Hill, New Jersey. The permutation group G generated by $(1, 2, \ldots, m)$ and $(1, 2, \ldots, n)$, where $1 < m < n$, is the alternating group A_n on n letters if $m \equiv n \equiv 1 \pmod 2$ and the entire symmetric group S_n otherwise.

We first show that $A_n \subseteq G$ in all cases. Let $C_m = (1, 2, \ldots, m)$ and $C_n = (1, 2, \ldots, n)$, and let composition of group elements proceed from right to left so that, e.g. (12) (13) = (132). To show $A_n \subseteq G$ it is sufficient to exhibit in G all 3-cycles, and since $(ijk) = (1ij)(1jk)$ it is sufficient to produce all 3-cycles containing 1, i.e. $(1jk)$. We first exhibit a 3-cycle as a commutator:

$$C_m C_n C_m^{-1} C_n^{-1} = (1, 2, m+1)$$

where $1 < m < n$ is used in this calculation. We produce the other 3-cycles by repeated conjugations. First

$$C_n^j (1, 2, m+1) C_n^{-j} = (j+1, j+2, m+j+1)$$

where the symbols in this and all following 3-cycles are to be interpreted as least positive residues (mod n). Next, starting with $(1, 2, m+1)$ corresponding to $j = 1$ and proceeding inductively on j we construct $(1, k, m+1)$ for all k:

$$(1, j+2, m+1) = (j+1, j+2, m+j+1)(1, j+1, m+1)(j+1, j+2, m+j+1)^{-1}.$$

Finally $(1, j, m+1)\ (1, k, m+1)\ (1, j, m+1)^{-1} = (1, j, k)$ finishes the construction.

Now $G = S_n$ or A_n according as G contains an odd permutation or not. The condition for a k-cycle to be an odd permutation is that $k \equiv 0 \pmod 2$. Hence $G = A_n$ if and only if C_m and C_n are both even permutations. I.e., $m \equiv n \equiv 1 \pmod 2$, and $G = S_n$ otherwise.

Also solved by J. L. Brenner, J. L. Bryant & R. W. Gilmer, C. V. Heuer, Michael Hoffman, I. M. Isaacs, A. A. Jagers (Netherlands), Robert Maas, Ralph Neuhaus, A. Sinkov, John H. Smith, Alena Vencovská & Aleš Drápal (Czechoslovakia), and the proposer.

Note. Eric Rosenthal points out that an APL computer program to calculate order of a subgroup of a symmetric group given the generators of the subgroup appears in Charles C. Sims, The influence of computers on algebra, Proceedings of symposia in applied mathematics, volume 20, 1974, *The Influence of Computing on Mathematical Research and Education*, edited by Joseph P. LaSalle, pp. 13–30.

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN
with the assistance of the mathematics departments of St. Olaf and Carleton Colleges
COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, CARLETON COLLEGE

We invite readers to submit reviews of significant recent college-level mathematics books. We especially encourage reviews based on classroom use, or comparative reviews of several related books. Reviews should ordinarily not exceed two pages (per book) typed double spaced. Manuscripts of reviews as well as books submitted for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057.

Elementary Number Theory: A Computer Approach. By Allan M. Kirch. Intext, New York, 1974. xi + 339 pp. $11.75. (Telegraphic Review, October 1974.)

This book was my choice for a text in a one-semester reading course. During that semester we covered all but the last section (Palindromic numbers) and the sections which required arithmetic using numbers with more than seven significant decimal digits. All the programs in the book are written in Fortran; we used Basic.

In my opinion this book provides an interesting mixture of theory and application both in the expository text and in the exercises. Among the topics examined are Diophantine equations, linear and non-linear, congruences, number bases, primes, Euler's ϕ function, perfect numbers, primitive roots and indices, quadratic residues, and continued fractions. For me the computer aspects made the subject more rewarding and student reaction to this aspect was favorable.

There were some difficulties. Two theorems on quadratic residues contained mistakes that were readily apparent. In my judgment the random number discussion would have been more valuable if the production of random numbers in other distributions, using those from the uniform distribution of random numbers, had been discussed, as all computer languages have a random number function. I also wish that not all the answers to the exercises had been given. Lastly, I believe that students without previous experience with computers would find the struggle of learning a language and number theory too demanding to enable them to progress very far in number theory.

In conclusion, my opinion is that this would make an excellent addition to a college library and, for classes with computing facilities available, this book has my endorsement.

GEORGE H. DUBAY, University of St. Thomas

Vector Calculus. By Jerrold E. Marsden and Anthony J. Tromba. W. H. Freeman, San Francisco, 1976. xiv + 449 pp. $16.00. (Telegraphic Review, March 1976.)

It has not always been easy to select a text for our sophomore-level course in several-variable calculus. Some books are too theoretical, and some are old-fashioned. Some books cover too many topics, and others don't go far enough. In addition, the structure of our program does not allow us to assume that students have had any linear algebra. Yet we want to offer a course that takes advantage of the notation and techniques of linear algebra. In the fall of 1975 we again selected Seeley's text *Calculus of Several Variables*, Scott, Foresman and Co., 1970. We were quite satisfied with Seeley, although the book is too theoretical for our average student. In January 1976 the text under review appeared in our mailboxes. After elaborate discussions with our students, we decided to change texts right in the middle of the term! *Vector Calculus* is designed for a one-semester or a two-quarter course, but we used it for the rest of the year since we naturally lost some time changing texts.

We were not disappointed. Marsden and Tromba have written a well-balanced text that will serve as an excellent bridge between the standard calculus course and the more theoretical upper division mathematics courses. Chapter 1 contains an introduction to the linear algebra needed in the rest of the

book; emphasis is placed on 2 and 3 dimensions. Standard differentiation topics, including gradients, directional derivatives and partial derivatives, are covered in Chapter 2; careful proofs are collected together in an optional section. Chapter 3 is a short chapter on vector-valued functions and vector fields, including the divergence and curl. Another short chapter, the fourth, deals with maximum and minimum problems and includes Taylor's theorem and Lagrange multipliers.

The second half of the book is devoted to integration. Chapter 5 contains a very nice treatment of multiple integrals. Results are clearly and correctly stated even when no proofs are provided. For example, Fubini's theorem in $\mathbf{R}^2$ is given for a bounded function with domain a rectangle whose set of discontinuities has area zero. This, in turn, is used to justify the definition of double integrals over more general regions. The last chapter, Chapter 7, ties the differentiation and integration theory nicely together and includes Green's theorem, Stokes' theorem and Gauss' theorem. This chapter is very comprehensible because it is preceded by a long chapter, Chapter 6, in which line and surface integrals of both scalar and vector valued functions are carefully developed. The connections with physics are made clear throughout, although physics is not a prerequisite. Unfortunately, instructors who want examples and motivation from other disciplines will have to supply them themselves.

The text is very student-oriented. Answers are given for about half of the problems. The book is very attractive with excellent pictures, including some computer-generated graphs. The text is almost completely free of errors and misprints and very few problems have wrong answers. These features make the book a pleasure to deal with, but the really fine feature is the authors' ability to introduce difficult concepts without overwhelming the reader with secondary matters.

RICHARD M. KOCH and KENNETH A. ROSS, University of Oregon

TELEGRAPHIC REVIEWS

Telegraphic reviews are designed to give prompt notice of new books with sufficient information to assist our readers in deciding whether to order an examination copy or to suggest library purchase. Possible uses are indicated as follows:

T = textbook P = professional reading
S = supplementary reading L = undergraduate library purchase
13 to 18 = freshman to second year graduate level usage
1 to 4 = appropriate time in semesters to cover text

Asterisks (*) or question marks (?) denote special positive or negative emphasis, respectively. Publishers are denoted by standard abbreviations; complete addresses may be found in *Books in Print*.

GENERAL, T(13: 2), *Topics in Mathematics*. Donald R. Burleson. P-H, 1977, viii + 520 pp, $12.95. Topics include sets and logic, number theory, matrices, statistics. Each section includes "Things to Ponder": questions and problems to stimulate discussion. Many examples; not enough exercises. JG

PRECALCULUS, T(13: 1), *Modern College Algebra and Trigonometry, Third Edition*. Edwin F. Beckenbach, Irving Drooyan. Wadsworth, 1977, 486 pp, $13.95. Revision of the second edition (TR, August-September 1972) includes an extension of chapters on basic algebra concepts. Also major reorganization of material on functions. LLK

PRECALCULUS, T(13), *Plane Trigonometry*. Allyn J. Washington, Carolyn E. Edmond. Cummings, 1977, 372 pp, $10.95. After a brief introduction to functions, discusses trigonometric functions (without wrapping function), identities, logarithms, and complex numbers. SG

FOUNDATIONS, T(16-17: 1), S, L, *Introduction to Computability*. Fred Hennie. A-W, 1977, ix + 374 pp, $20.95. Presents some models of algorithmic processes and characterizes functions that can be evaluated algorithmically. Implication of families of algorithms for the capabilities and limitations of programming languages. Turing machines, recursive functions, computability, and decidability. Chapter problem sets. References. Index. RJA

FOUNDATIONS, S(15), P, *The Illusory Infinite--A Theology of Mathematics*. J. Fang. Paideia Pr, 1976, 351 pp, $9.80 (P). The author attempts to focus on the paradoxical and theological nature of "The Transfinite Arithmetic." He succeeds in his goal of being "thoroughly and openly long-winded in [his] efforts to present some primitive facts in mathematics as fully and clearly as possible." CEC

COMBINATORICS, P, *Polynômes arithmétiques et Méthode des Polyèdres en Combinatoire*. E. Ehrhart. Int. Ser. Num. Math., V. 35. Birkhäuser, 1977, 165 pp, sFr. 32 (P). A study of polynomials associated with periodic sequences of rational numbers arising in combinatorics and number theory. SG

NUMBER THEORY, P, *Introduction to Modular Forms*. Serge Lang. Grund. math. Wissenschaften, B. 222. Springer-Verlag, 1976, ix + 261 pp, $22.20. Intended as an introduction to the subject, but written with an eye toward the relationship between modular forms and representation of Galois groups of number fields. Topics include Hecke operators, Petersson scalar product, cusp form on $SL_2(Z)$, modular forms for congruence subgroups, congruence properties, p-adic distributions. SG

NUMBER THEORY, S(17), P, *Introduction to Modular Forms*. A. Robert. Pure and Appl. Math., No. 45. Queen's U, 1976, 98 pp, (P). A set of lecture notes with the purpose of introducing graduate students to some modern aspects of the theory of one variable modular form. CEC

LINEAR ALGEBRA, T(14), *Elementary Linear Algebra, Second Edition*. Howard Anton. Wiley, 1977, xv + 351 pp, $13.95. A sophomore level introduction that does not presume calculus. Begins with systems of linear equations, determinants (of n x n matrices) and vectors in 2- and 3-space before considering Euclidean vector spaces, linear maps, eigenquestions and applications. Changes from first edition (TR, March 1973; ER, March 1974) include some rewriting and a separate paperback supplement with problems on applications. SG

ALGEBRA, T?(18: 1), S, P*, *The Theory of Unitary Group Representations*. George W. Mackey. U of Chicago Pr, 1976, x + 372 pp, $4.95 (P); $12. A series of lectures given by the author in 1955 along with an extensive appendix that sketches the developments which have taken place since that time. An exhaustive bibliography is also included. CEC

ALGEBRA, P, *Coherence and Non-Commutative Diagrams in Closed Categories*. Rodiani Voreadou. Memoirs No. 182. AMS, 1977, xvi + 93 pp, $7.20 (P). A procedure for testing (in a finite number of steps) whether or not a given diagram of natural transformations commutes in every closed category. PJM

ALGEBRA, P, *Enveloping Algebras*. Jacques Dixmier. Math. Lib., V. 14. North-Holland, 1977, xvi + 375 pp, $36.75. A well-motivated, well-written introduction to enveloping algebras. Covers basic theory of said algebras (two-sided ideals, centres, primitive ideals) and the connections between representations of enveloping algebras and representations of Lie algebras. SG

ALGEBRA, T(18), P, *Modular Representations of Finite Groups*. B.M. Puttaswamaiah, John D. Dixon. Pure and Appl. Math., V. 73. Acad Pr, 1977, xv + 242 pp, $23.50. A comprehensive introduction to modular representation theory. Provides necessary background from algebra and ordinary representation theory. Detailed consideration of blocks, indecomposable modules, and Brauer's main theorems. A few exercises included. A valuable reference. SG

CALCULUS, T(13:3), *Calculus and Analytic Geometry, Second Edition*. Sherman K. Stein. McGraw, 1977, xiv + 994 pp, $18.95; *Instructor's Manual*, 119 pp, (P). Changes from the first edition (TR, November 1973; ER, February 1976) include an expanded chapter on series, new material on determinants, and even more exercises and examples. As before, the chapter summaries are commendable. JG

CALCULUS, T(13: 1). *Technical Calculus.* Dale Ewen, Michael A. Topper. P-H, 1977, x + 374 pp, $12.95. Designed to provide calculus for students enrolled in an engineering technology program which requires practical calculus. Includes differentiation, integration, power series, some numerical methods and some differential equations. A standard cookbook approach. CEC

REAL ANALYSIS, T(17: 1, 2), S. *Real Variable and Integration with Historical Notes.* John J. Benedetto. Teubner, Stuttgart, 1976, 278 pp, $24 (P). Text for basic graduate course in real variable which emphasizes the notion of absolute continuity, a concept which unifies Lebesgue's dominated convergence theorem and the Radon-Nikodým theorem. Includes historical notes. Is a commercial for Vitali's contributions. Large problem sets. 129 references. Index of names, of terms. RBK

REAL ANALYSIS, T(17: 2). *Real Analysis.* Nicolas K. Artémiadis. So Ill U Pr, 1976, xii + 581 pp, $12.50. A very formal treatment of the first year graduate course. The book contains no illustrations at all. Topics: real numbers, Lebesgue integration, metric and L^p spaces, topological spaces, Banach and Hilbert spaces. Includes a wide selection of exercises. Students may find this to be dry reading. TAV

COMPLEX ANALYSIS, T(16), L. *Elements of Complex Analysis.* Jacob Sonnenschein, Simon Green. Dickenson, 1977, viii + 280 pp, $14.95. Aim is to develop basic theory as rigorously and quickly as possible. Properties of differentiable functions, holomorphic functions (Cauchy's theorem is true), analytic functions (power series representation), and conformal mappings are developed in separate chapters, but shown to be equivalent. RBK

COMPLEX ANALYSIS, P. *Riemann Surfaces and Generalized Theta Functions.* Robert C. Gunning. Ergebnisse der Math., V. 91. Springer-Verlag, 1976, xii + 165 pp, $19.70. Generalized theta functions are vector-valued functions which represent cross-sections of some complex vector bundles over complex tori associated with compact Riemann surfaces. RBK

DIFFERENTIAL EQUATIONS. *Solution of Boundary Value Problems by the Method of Integral Operators.* D.L. Colton. Pitman, 1976, 148 pp, $5.60 (P). Companion lecture notes to author's *Partial Differential Equations in the Complex Domain* which discussed integral operator methods only briefly. Develops details of interplay between analytic continuation and the approximation of solutions to second order partial differential equations. RBK

DIFFERENTIAL EQUATIONS, T(16: 1). *Operational Calculus.* Gregers Krabbe. Plenum, 1975, xvi + 349 pp, $8.95 (P). Paperback version of 1970 Springer edition. Similar to Mikusiński calculus, but uses an algebra isomorphic to Schwartz's convolution algebra rather than Mikusiński's field of convolution quotients. Examples, but no exercises. RBK

DIFFERENTIAL EQUATIONS, T(14-15). *Ordinary Differential Equations for Engineering and Science Students.* L.B. Jones. Bradford U Pr, 1976, ix + 220 pp, £3.25 (P). Intended for students of engineering and physical science. Covers first order equations, higher order equations with constant coefficients, numerical methods for first order to higher order equations, series solutions, Laplace transforms. Few exercises. SG

DIFFERENTIAL EQUATIONS, T(15-16), L*. *Ordinary and Delay Differential Equations.* R.D. Driver. Appl. Math. Sci., V. 20. Springer-Verlag, 1977, ix + 501 pp, $14.80 (P). The first half provides a fairly rigorous introduction to ordinary differential equations (uniqueness and Lipschitz conditions, with order equations, linear systems). The second half offers a well-motivated fairly rigorous introduction to delay differential equations (existence, linear delay differential systems, stability, autonomous systems). Many exercises. An interesting book. SG

FUNCTIONAL ANALYSIS, P. *Projection-Iterative Methods for Solution of Operator Equations.* N.S. Kurpel'. Trans. Math. Mono., V. 46. AMS, 1976, 196 pp, $24.80.

FUNCTIONAL ANALYSIS, P. *Bifurcation Theory for Fredholm Operators.* Jorge Ize. Memoirs No. 174. AMS, 1976, viii + 128 pp, $7.60 (P).

OPTIMIZATION, T(15-17: 1), S, L. *Mathematische Methoden des Operations Research.* Peter Kall. Teubner, Stuttgart, 1976, 176 pp, DM 22,80 (P). Mathematical methods of operations research. Linear optimization: examples of production, combination, transportation, network flow problems; linear programming techniques. Nonlinear optimization: convex sets; convex functions; convex programming; Kuhn-Tucker Theorem. Short introduction to dynamic optimization. Bibliography. Index. RJA

ANALYSIS, T(18), P. *Introduction to Ergodic Theory.* Ya. G. Sinai. Trans: V. Scheffer. Princeton U Pr, 1976, 144 pp, $6 (P). A brief, attractive introduction concentrating on "the study of statistical properties of groups of motion of non-random objects." Among the main topics are invariant measures under groups of transformations, dynamical systems, billiard balls, and entropy. SG

TOPOLOGY, P. *Characteristic Classes of Foliations.* H.V. Pittie. Pitman, 1976, 107 pp, $8.75 (P). A foliation of a manifold is a decomposition into disjoint submanifolds which glue together smoothly. Characteristic classes of foliations are (homotopy) invariants living in homology groups. This monograph, based on the author's lectures at Warwick University in 1975, surveys recent results on characteristic classes of foliations. PJM

TOPOLOGY, P. *Lecture Notes in Mathematics-542: Čech and Steenrod Homotopy Theories with Applications to Geometric Topology.* David A. Edwards, Harold M. Hastings. Springer-Verlag, 1976, vii + 296 pp, $11.50 (P). Homotopy based on limits of spaces applied to problems in piecewise linear topology. PJM

TOPOLOGY, P. *Hopf Spaces.* Alexander Zabrodsky. Math. Stud., V. 22. North-Holland, 1976, x + 223 pp, $18.50 (P). A Hopf space (or H-space) is a topological space with a multiplication. This monograph collects some recent results on H-spaces, including the author's "mixing homotopy types" method of constructing new ones. PJM

TOPOLOGY, P. *Topology: Proceedings of the Memphis State University Conference.* Ed: Stanley P. Franklin, Barbara V. Smith Thomas. Lect. Notes in Pure and Appl. Math., V. 24. Dekker, 1976, xii + 296 pp, $24.50 (P). Five of six invited addresses and twenty-one of thirty-six contributed papers from the March 1975 ninth annual spring topology conference which was mostly devoted to point set topology. JAS

TOPOLOGY, T(15-17: 1), S, L. *Anschauliche Topologie.* Kurt Peter Müller, Heinrich Wölpert. Teubner, Stuttgart, 1976, 168 pp, DM 18,80 (P). Intuitive topology made precise. Basics of point-set topology. Graphs in the plans and maps. Lines and knots. Topology in three-dimensional space: surfaces. Many geometric diagrams and tables. Exercises and answers. References. Index. RJA

STATISTICS, T(18: 1). *Introduction to Statistical Time Series.* Wayne A. Fuller. Wiley, 1976, ix + 470 pp, $24.95. Text for graduate course in time series. Presumes graduate courses in statistical theory and linear regression analysis. Includes results from Fourier analysis, large sample statistics, and difference equations. Theorem-proof format. Exercises use moderate size data sets. 200 references. Index. RBK

STATISTICS, T(18: 1), P. *Lecture Notes in Mathematics-520: Techniques of Multivariate Calculation.* Roger H. Farrell. Springer-Verlag, 1976, x + 337 pp, $12.30 (P). Methods for the calculation of multivariate densities which have been used in the literature, but not in books: transforms, decomposition of manifolds, factorization of invariant measures, and "random variable" techniques. RBK

STATISTICS, T(14: 1). *Statistics and Experimental Design for Behavioral and Biological Researchers: An Introduction.* Victor H. Denenberg. Wiley, 1976, viii + 344 pp, $17.50. Statistical methods for animal researchers in biology and psychology. No probability theory. Practical advice on applying methods. Few problems, but solutions to all. RBK

STATISTICS, T*(13: 1). *Statistics: Basic Techniques for Solving Applied Problems.* Stephen A. Book. McGraw, 1977, xii + 511 pp, $14.95. A text at precalculus level. Effective use of set diagrams, tree diagrams. Summary and discussion, bibliography, supplementary exercises for each chapter. Unusual variety of examples, problems. Glossary of formulas; answers to odd problems. RBK

STATISTICS, T*(1, 2), L. *Probability and Statistical Inference.* Robert V. Hogg, Elliott A. Tanis. Macmillan, 1977, ix + 450 pp, $14.95. An attractive text for a 3-6 hour post-calculus course in statistics. Similar in spirit but at a lower level than Hogg and Craig, *Introduction to Mathematical Statistics.* An unusual feature is the use of order statistics to make the first inference about percentiles and means. RBK

STATISTICS, T*(16-17: 2), L. *An Introduction to Probability Theory and Mathematical Statistics.* V.K. Rohatgi. Wiley, 1976, xiv + 684 pp, $22.95. From the preface: "...this is a mathematics text and not a 'cookbook'. It should not be used for service courses." Indeed this is a carefully written, solid introduction to probability and statistics through ANOVA with additional material on non-parametric and sequential inference. Presumes, and uses, advanced calculus, and linear algebra. A well conceived and executed text.

STATISTICS, T(16: 2). *Fourier Analysis of Time Series: An Introduction.* Peter Bloomfield. Wiley, 1976, xiii + 258 pp, $18.95. Intended to serve as a text for a course in time series as well as a reference for workers in fields in which time series arise. In deference to the second group the statistical elegance is replaced by numerous examples. Computational aspects are based upon the Fast Fourier transform, including Fortran programs. TAV

COMPUTER SCIENCE, S(17-18), P. *Tutorial on Software Design Techniques.* Peter Freeman, Anthony I. Wasserman. IEEE Comp. Soc., 1976, vi + 277 pp, $12 (P). An outgrowth of a one-day tutorial presented at the Second International Conference on Software Engineering. Contains eighteen papers on software design. Each group of papers is preceded by an introductory overview of the main topic under discussion. These overviews attempt to provide a general framework into which the papers fit. Annotated bibliography. RJA

COMPUTER SCIENCE, S(15-17). *Tutorial on Designing with Microprocessors.* Tilak Agerwala, Gerald Masson, Roger Westgate. IEEE Comp. Soc., 1976, 168 pp, $10 (P). An easy-to-read introductory overview of the design and architecture of currently available microprocessors. Saturated with design diagrams, logic diagrams, and circuit diagrams. Contains a comparison of maxis, minis, micros, and hardwired logic. Chapter references. RJA

COMPUTER SCIENCE. *Debugging System 360/370 Programs Using OS and VS Storage Dumps.* Daniel H. Rindfleisch. P-H, 1976, vii + 262 pp, $15.95.

COMPUTER SCIENCE, T(13-16), S, L. *Algorithms + Data Structures = Programs.* Niklaus Wirth. P-H, 1976, xvii + 366 pp, $15.50. Uses the syntax of Pascal to describe data structures and their algorithms. Static and dynamic structures. Internal and external sorting. Chapter on examples using recursive algorithms. Treats parsing and the construction of a simple compiler as an application of recursive techniques. Exercises. Chapter references. Appendices. Subject index. Program index. RJA

Reviewers Whose Initials Appear Above

Richard J. Allen, St. Olaf; Clifton E. Corzatt, St. Olaf; Jennifer Galovich, St. Olaf; Steven Galovich, Carleton; Lorraine L. Keller, St. Olaf; Roger B. Kirchner, Carleton; Pierre J. Malraison, Carleton; J. Arthur Seebach, Jr., St. Olaf; T.A. Vessey, St. Olaf.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, NW, Washington, D.C. 20036. Items must be submitted at least five months before publication can take place.

PERSONAL ITEMS

Professor Billy E. Rhoades, Indiana University, represented the Association at the inauguration of Dr. Samuel Foster Hulbert as President of Rose-Hulman Institute of Technology on November 14, 1976.

Clemson University: Associate Professors J. D. Fulton and D. R. LaTorre have been promoted to Professors; Assistant Professor Sue K. Dunkle retired in May 1976 with the title of Assistant Professor Emeritus.

Emory University: Professor Trevor Evans has been appointed Chairman of the Mathematics Department; Dr. J. K. Doyle, Syracuse University, has been appointed Assistant Professor; Dr. D. J. Pokrass, University of Missouri, Columbia, has been appointed Assistant Professor.

J. Sargeant Reynolds Community College: Dr. M. P. Eisner, Ball State University, has been appointed Assistant Professor; Assistant Professor and former Mathematics Program Head E. D. Bender has been appointed Acting Division Chairman of Natural Sciences and Mathematics.

Pan American University: Assistant Professor Step 2 T. F. McCabe has been promoted to Assistant Professor Step 3; Instructor R. A. Close has been promoted to Assistant Professor Step 1.

University of California, Davis: Professors H. L. Alder and G. D. Chakerian were recipients of 1976 Distinguished Teaching Awards.

University of Colorado, Denver: Assistant Professor Zenas Hartvigson has been promoted to Associate Professor; Associate Professor Collin Hightower has been promoted to Professor.

University of Texas at San Antonio: Dr. Manuel Berriozabal, University of New Orleans, has been appointed Professor; Dr. Joseph Valentine, Utah State University, has been appointed Professor.

Dr. L. A. Aroian, Professor Emeritus of Administration and Management at Union College, has been appointed Visiting Professor of Mathematics at Western Illinois University for the period September 1976 to June 1977.

Associate Professor Irving Bentsen, Chairman of the Department of Mathematics at Hobart and William Smith Colleges, has been promoted to Professor.

Assistant Professor D. C. Cathcart, Salisbury State College, has been promoted to Associate Professor.

Assistant Professor Hans Herda, Boston State College, has been promoted to Associate Professor.

Instructor J. M. Lamb, Tri-County Technical College, has been appointed Chairman of the Department of Mathematics.

Dr. R. J. Larsen, SUNY at Binghamton, has been appointed Visiting Associate Professor at SUNY at Albany.

Mrs. Myra J. Reed has been appointed Visiting Assistant Professor at Vassar College.

Dr. V. R. R. Uppuluri, on leave of absence from Union Carbide-Nuclear Division, Oak Ridge, has been appointed Visiting Professor at the University of California, Santa Barbara, for the academic year 1976–77.

Associate Professor Benjamin Wells, University of Hawaii, has been promoted to Professor.

Professor Emeritus John L. Barnes, UCLA, died on October 1, 1976, at the age of 69. He was a member of the Association for forty-six years.

Mr. Kenneth C. Cartwright, Glendale, California, died on September 18, 1976. He was a member of the Association for nineteen years.

Professor Emeritus Carl J. Coe, University of Michigan, died on August 24, 1976, at the age of 95. He was a member of the Association for fifty-three years.

Professor R. C. Morrow, United States Naval Academy, died on August 18, 1976, at the age of 62. He was a member of the Association for twenty-one years.

Dr. Christos D. Papakyriakopoulos, Princeton University, died on June 29, 1976, at the age of 62. He was a member of the Association for nineteen years.

A NEW STUDY OF THE NON-ACADEMIC JOB MARKET— A CALL FOR PARTICIPATION

Most projections of supply and demand for Ph.D.'s in science and engineering over the next decade or two estimate that the number of new doctoral recipients will provide about one-third more Ph.D.'s by 1985 than requred for academic and R & D openings.

To address some of the problems inherent in this projected surplus, the Higher Education Research Institute (HERI) in Los Angeles has launched a project, funded by the Ford and National Science Foundations, to identify nontraditional job markets for academics and determine entry points for new Ph.D. recipients.

Since information about the migration of Ph.D.'s into and out of academic employment and within or between academic institutions and business is the weakest link in the knowledge of manpower flows, HERI will first endeavor to refine projections of job availability by collecting data on mobility of doctorate holders. The HERI study will determine the most fruitful nontraditional markets and whether they represent "enrichment" or serious "underutilization" or underemployment of Ph.D.'s. It will also suggest new areas for productive employment.

HERI will survey those in nontraditional jobs to determine why doctorate holders take such employment and to evaluate job satisfaction and other career outcomes. The study will compare backgrounds, training, and attitudes of those in traditional and nontraditional careers.

To analyze the reasons behind the moves of mobile doctorate holders, the HERI study will ask such questions as: Were Ph.D.'s attracted to new jobs by good offers, or were they forced from previous jobs? Are those who change jobs in the current market happier in their new jobs? How could intersectoral moves be encouraged or reduced in the future?

The study will sample 15,000 doctorate holders from 14 science (physical, biological, and social) and engineering fields who have moved out of or into nonacademic jobs in the last three years. The nontraditional employment areas into which senior personnel have moved represent potential new markets for science and engineering doctorates.

HERI will also survey Ph.D.'s who have moved between academic institutions, from academic or research to administrative positions within a college, university, or research organization, or within or between business firms. The study will explore the career patterns, characteristics, and routes of access of the highly mobile Ph.D.'s.

The Study has three major objectives:

1. To identify nontraditional job markets outside academe which offer satisfaction and utilize science and engineering backgrounds. The study will evaluate the extent to which these nontraditional markets can absorb the expected surplus of Ph.D.'s and the "push and pull" factors experienced by movers.

2. To determine potential entry points for new degree recipients by looking at careers of current doctorate holders.

3. To collect data on the occupational mobility of science and engineering Ph.D.'s in nontraditional fields, with particular attention to the "fine" fields, those specializations within major areas.

HERI will identify the sample population through professional science and engineering associations, the academic departments of 160 colleges and universities, and 200 major U.S. corporations listed in the *Fortune* 500. It will solicit participation through advertisements in several major national publications.

Requirements for participation in the study include a Ph.D. in any of the following fields: economics, mathematics, chemistry, physics, sociology, biology, political science, anthropology, psychology, civil engineering, electrical engineering, mechanical engineering or zoology.

HERI is seeking the participation of Ph.D.'s who hold a nontraditional or unusual job outside the academic or traditional research areas, or who have changed employers or job functions within the past three years. Those who would like to participate in the one-year study should write to: Higher Education Research Institute, 924 Westwood Boulevard/Suite 850, Los Angeles, California 90024.

(*The results of the HERI study may benefit mathematicians greatly in the near future. We urge our readers who hold Ph.D.'s in science or engineering and who are employed in non-academic positions to participate in this study. Editor.*)

SEMINAR SPONSORED BY THE NORTH CENTRAL SECTION

The North Central Section of the MAA will sponsor a seminar on Applications of Mathematics in Modeling Theory from June 20, 1977, to June 24, 1977, at Bemidji State University, Bemidji, Minnesota. Applications presented will in general be of such a nature that they could be used in the teaching of undergraduate classes. Cost for room and board on the campus together with registration will be approximately seventy dollars. Northern Minnesota has many beautiful resort areas so if you wish to bring the family and camp while attending the seminar,

please contact Dr. Clayton Knoshaug of Bemidji State University. If you wish to attend and/or to present a paper, please contact Dr. Gerald Bergum, South Dakota State University, Brookings, SD 57006. Two hours of credit can be earned. We have already arranged to have speakers in attendance from St. Paul Fire and Marine Insurance, Minnesota Mutual Life Insurance, Contral Data Corporation, Honeywell Corporation, and Mayo Clinic.

AN APPEAL

Universities of developing countries are generally able to subscribe to only a very limited number of scientific journals. Steeply rising prices often compel them to reduce their subscriptions from one year to the next. Many libraries and science departments of developing countries will be grateful to receive donations of back-volumes of journals that individuals or institutions may like to dispense with. The International Centre for Theoretical Physics, Trieste, is willing to help establish contacts among possible donors and recipients, as far as journals of physics/mathematics are concerned. Please address inquiries and offers of donation to: Professor L. Fonda, International Centre for Theoretical Physics, P.O. Box 586, Trieste, I–34100 Italy.

CARL B. ALLENDOERFER TEACHING FELLOWSHIPS

The Mathematics Department of the University of Washington, Seattle, has established two Carl B. Allendoerfer Teaching Fellowships for first-year graduate students.

UNPAID RESEARCH APPOINTMENTS AT THE UNIVERSITY OF NEW HAVEN

The University of New Haven, West Haven, Connecticut, announces a proposed informal agreement which will make university facilities available to professional mathematicians who are in the area and who have no academic connection, or whose "first home" has inadequate facilities.

This informal agreement, providing a "second home", is an unpaid research appointment.

Anyone interested in making such an arrangement should contact Bertram Ross, Professor of Mathematics, University of New Haven, New Haven, CT 06505.

FIFTH ANNUAL MATHEMATICS AND STATISTICS CONFERENCE MIAMI UNIVERSITY, OXFORD, OHIO

The Fifth Annual Mathematics and Statistics Conference at Miami University, Oxford, Ohio, will be held September 30–October 1, 1977. The theme is "Number Theory—Pure and Simple", and Professor Ivan Niven of the University of Oregon will be a featured speaker. On Friday the talks will be directed at college teachers and researchers; on Saturday they will be directed at high school teachers and students at all levels. There will be sessions for contributed papers, and abstracts should be sent to Dr. Stanley Payne, Department of Mathematics and Statistics, Miami University, Oxford, Ohio 45056. The deadline for abstracts is July 1, 1977. (Late abstracts may be considered.) Information concerning preregistration, housing, etc., may also be obtained from the above address.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

THE SIXTIETH ANNUAL MEETING OF THE ASSOCIATION

The Sixtieth Annual Meeting of the Mathematical Association of America was held at the Chase-Park Plaza Hotel, St. Louis, Missouri, from Saturday to Monday, January 29–31, 1977, in conjunction with meetings of the American Mathematical Society, Association for Symbolic Logic, Association for Women in Mathematics, Conference Board of the Mathematical Sciences, Mathematicians Action Group, National Association of Mathematicians, and the National Council of Teachers of Mathematics. The Saturday and Sunday morning sessions of the Association were joint meetings with the National Council of Teachers of Mathematics; the Saturday evening session was a joint meeting with the American Mathematical Society. There were registered 2993 persons, including 1741 members of the Association

Sessions were held on Saturday morning, Saturday evening, Sunday morning, Monday morning and Monday afternoon in the Khorassan Ballroom. Presiding officer at the lecture by Dr. R. H. Hudgin was Dr. Gordon Raisbeck; at the lecture by Professor K. I. Gross, Professor E. N. Wilson; at the lecture by Professor H. S. Wilf, Dr. Marjorie L. Stein; at the lecture by Professor Mary Ellen Rudin, Professor R. H. McDowell; at the lecture by Professor Leonard Gillman, Professor D. P. Roselle; at the lecture by Dr. B. O. Koopman, Professor Leonard Gillman.

The Program Committee consisted of G. L. Weiss, Chairman; D. J. Albers, R. W. Freese, Deborah T. Haimo, Nancy T. Holder, R. H. McDowell, D. P. Prekeges, Gordon Raisbeck, and Marjorie L. Stein.

FIRST SESSION OF THE ASSOCIATION

Joint with the National Council of Teachers of Mathematics

Panel Discussion: *The Third International Congress on Mathematics Education.*

A panel discussion with Professor Shirley A. Hill, University of Missouri, Kansas City, Professor P. J. Hilton, Battelle Memorial Institute and Case Western Reserve University, Professor Jeremy Kilpatrick, University of Georgia, Professor Seymour Schuster, Carleton College, moderated by Professor G. S. Young, University of Rochester.

Professor Hilton gave attention to the content of the five main addresses at the Karlsruhe International Congress on Mathematics Education (August, 1976). The objectives and achievements of Section B3, in which the speaker participated, were also discussed. This Section sought to answer the question, 'Why do we teach mathematics?' The consensus arrived at in this Section is to be incorporated in the UNESCO publication, "New Trends in Mathematics Education," which is expected to have a significant influence in the 'third world'.

Professor Kilpatrick said that a theme common to several reports at ICME 3 was that mathematics educators need to give more attention to the context in which teaching and learning occur. This context includes the economic, social, political, and cultural forces operating inside and outside classrooms to influence instruction. In many countries there have been recent assessments of educational priorities, partly but not entirely, because of rising costs. Mathematics educators trying to improve school instruction — whether by creating curriculum materials, doing research on learning, or educating prospective teachers — have had to respond to shifting demands and new educational climates.

Professor Schuster said that the program of the International Congress held at Karlsruhe was formally divided into four parts. Part I, entitled Main Papers, consisted of one-hour plenary sessions at which five internationally known scholars presented papers intended to be of interest to all participants. Part II, entitled Sections, was the main focus of ICME III. There were 13 sections, each meeting at least twice, devoted to discussions of survey-trend reports prepared by Reporters, who were specialists in the areas of concern. Part III, The Poster Sessions, consisted of the presentation and informal discussion of papers that were, literally, posted in hallways. Part IV, which concluded the Congress, consisted of a panel discussion on the long range implications of computers and calculators for mathematics education. Among the critical observations made on the format and organization of the Congress were that the main work of ICME, namely the Sections, were too much in conflict with one another; further, they gave little opportunity for participation by non-panelists.

Compensated Imaging-From Fantasy to Reality, by Dr. R. H. Hudgin, Electro-optics Division, Itek Corporation.

Dr. Hudgin reported that correcting for a turbulent atmosphere has been a dream of astronomers since Galileo and has recently become a reality at Itek Corporation. One of the critical technologies was an actively deformable mirror which could be distorted rapidly and precisely to correct for aberrations in the incident light. The analysis and development of this device was presented, and a movie demonstrating the world's first atmospheric compensation system was shown.

SECOND SESSION OF THE ASSOCIATION

Joint with the American Mathematical Society

Panel Discussion: *China Report*

A panel discussion with Professor V. L. Klee, Jr., University of Washington, Professor J. J. Kohn, Princeton University, MAA President H. O. Pollak, Bell Telephone Laboratories, moderated by Professor Saunders Mac Lane, University of Chicago.

This was a report on a visit to the People's Republic of China in May, 1976, by a delegation of pure and applied mathematicians sponsored by the Committee of the National Academy of Sciences on the Scholarly Communication with the People's Republic of China. Each of the four members of the panel presented certain aspects of the delegation's observations on mathematics in China. Victor Klee discussed Chinese work in operations research and in classical applied mathematics. He observed that Chinese activity in operations research, especially in optimal search, PERT, and other methods, is very active. J. J. Kohn spoke about the quality of Chinese mathematical activity and the effects of political restraints. Henry Pollak summarized the observations of the delegation on educational policy and activity in China, both in the middle school and in the university. Saunders Mac Lane summarized some of the research results recently obtained by Chinese mathematicians, especially those at the Mathematical Institute in Peking.

AMS-MAA Concert by Professor Howard Osborn and Ms. Kiyoko Takeuti.

The following program for piano and violin was presented:

Sonata in b flat, K454 (1784)	Mozart
Sonata in c minor, Op. 30, No. 2 (1802)	Beethoven
INTERMISSION	
Sonata in a minor, Op. 108 (1888)	Brahms

THIRD SESSION OF THE ASSOCIATION

Panel Discussion: *Decline in the Preparation of Students for Collegiate Mathematics*

A panel discussion with Professor G. L. Alexanderson, University of Santa Clara, Mr. D. R. Johnson, Nicolet High School, Milwaukee, Wisconsin, Ms. Katherine P. Layton, Beverly Hills High School, Beverly Hills, California, moderated by MAA President-Elect H. L. Alder, University of California, Davis.

The moderator announced that, a year ago, the MAA and NCTM jointly appointed a Committee on the Reported Decline in the Preparation of Students for Collegiate Mathematics. This action was motivated by the widely reported decline in the average Scholastic Aptitude Test scores in mathematics which have dropped every year from 1962 to 1975.

This joint Committee consists of three high school teachers of mathematics, one of whom is a mathematics supervisor, one two-year college teacher, and three college or university teachers.

In accordance with the charge given to the Committee, it has prepared a Statement listing priorities for mathematical topics and concepts for entering college students and agreed on eleven recommendations designed to reverse the decline. Both the Statement and the eleven recommendations have been submitted to the NCTM Board of Directors and the MAA Board of Governors. When approved, these will be widely distributed.

The moderator expressed his firm conviction that if the Statement and the eleven recommendations are widely followed in our schools, the result is bound to be a significant improvement in the preparation of students for collegiate mathematics. He then introduced three members of the Committee who reported on various aspects of the Committee's findings and recommendations.

Professor Alexanderson said that the Committee on the Reported Decline in the Preparation of Students for Collegiate Mathematics examined available reports and surveys on the reported decline in preparation of students and held meetings with teachers in various parts of the country. Reports of these studies and meetings were given and several proposed causes for the decline were examined. The statement on the preparation needed by students planning to take collegiate level courses in mathematics was presented.

Mr. Johnson announced the eleven recommendations (with accompanying rationale) of the Committee on the Reported Decline in the Preparation of Students for Collegiate Mathematics.

Ms. Layton recalled that in 1975 the College Entrance Examination Board appointed a Panel on the Decline in the Scholastic Aptitude Test Scores. The median score on the SAT-Math was 502 in 1963 (highest possible, 800); in 1976, it was 472. Why? The report of the Panel, which addresses this question, will be published in Spring, 1977. Possible causes lie with the test, population sitting for the test, and changes in society. At the request of the Panel, studies have been undertaken; for example, "Relevance of SAT to High School Curricula," "Television and Test Scores," "Some Trends in the Entry to Higher Education," "Review of Zajonc Hypothesis," and "Valedictorian Study."

Panel Discussion: *How to Cope with MAA-M*ath *A*voidance and *A*nxiety.

A panel discussion with Professor Lenore Blum, Mills College, Professor Karel DeLeeuw, Stanford University, Preceptor Deborah Hughes Hallett, Harvard University, Professor R. A. Rosenbaum, Wesleyan University, moderated by Professor Alice T. Schafer, Wellesley College.

The moderator noted that many women enter college (1) having been discouraged from taking more mathematics in high school than is necessary for college admission; (2) with fears, anxieties and/or dislikes toward mathematics; and (3) with little or no knowledge that mathematics is increasingly being used in a wide variety of fields and that, consequently, (4) without some knowledge of mathematics their career options are severely limited. At Wellesley a one semester course, A Discovery Course in Mathematics and its Applications, is being designed for such people. The course (1) emphasizes the usefulness of mathematics, particularly in fields other than the physical sciences; (2) the atmosphere in the class is a nonpressuring one with a teaching method of discussion by students and faculty on an equal basis; (3) in addition to two regular 70 minute periods per week, there is a weekly two-hour session where students and faculty work together; and (4) there is a lecture series by women who are not mathematicians but who use mathematics in their careers. In addition to an evaluation of the course on mathematical grounds, two psychologists are making an evaluation on psychological grounds: are mathematical anxieties being eradicated or lessened?; is any insight being gained into the determination of the causes of those anxieties?

Professor DeLeeuw reported that he and Phil Faillace have worked to create at Stanford a learning environment for students weak in mathematics. They have given parallel courses for students and tutors. The student course is flexible, emphasizing analytical skills and visualization. The main discovery has been that the university has a vast number of students who are eager to work as teachers. And that they can be extremely effective working with students whose main problem is not with the subject matter but with their attitude toward the subject matter.

Preceptor Hughes Hallet said that the success of a remedial math course depends critically on the attitudes of its teachers. Of crucial importance is their belief that the students can-and *will*-master the material, and a corresponding willingness to "reach out" to the students. If this means calling students who have temporarily given up, or taking the time to talk to them about their anxieties, then so be it — this is part of teaching math at this level. The truth of the matter is that students get more from our faith in their ability to learn math than from our best explanation or cleverest innovation, she said.

Professor Rosenbaum said the following were the noteworthy features of the Wesleyan experience:

(a) When a pre-calculus course was first offered for credit, enrollment was much greater than expected, and large enrollments have continued;

(b) The high level of apprehension of some students seems to be alleviated by their dealing with a mathematics "clinic," rather than a math "department";

(c) Counselling in the "clinic," including group discussions of attitudes about mathematics, seems to be of help to some of the most anxious students.

Professor Blum stated that the major premise is that the best way of overcoming math avoidance is by having students experience success by really "doing it." The assumption is that attitudes can change rapidly and that several significant positive experiences can make the difference. The Mills College program, by providing a support structure of peer-taught workshops for entry level courses, is designed to get students into a substantive mathematics curriculum quickly; it also provides students (even some of the most math anxious) opportunities to become actively involved in (peer) teaching and in (industrial) research projects. During the three years since the program began, the number of students enrolled in pre-calculus has tripled and the total number of enrollments in regular math and computer science courses (not including workshop students) has nearly doubled.

FOURTH SESSION OF THE ASSOCIATION

On the Evolution of Noncommutative Harmonic Analysis, by Professor K. I. Gross, University of North Carolina, Chapel Hill.

Professor Gross devoted his lecture to a quasi-historical, elementary account of the development of the subject

of non-commutative harmonic analysis. He traced the origins in and interconnections with algebra, analysis, geometry, and mathematical physics.

Trends in Undergraduate Mathematics: The CBMS Survey, by Professor D. J. Albers, Menlo College, and Professor J. W. Jewett, Oklahoma State University.

The report, "Undergraduate Mathematical Sciences in Universities, Four-Year Colleges, and Two-Year Colleges 1975–76" is available at $4 postpaid from CBMS, 2100 Pennsylvania Avenue, N.W., Washington, D.C. 20037.

Professor Albers reported that, from 1970 to 1975, in two-year colleges:

1. Mathematics enrollments increased by 50%. Two-year colleges now account for 37% of *all* undergraduate mathematical science enrollments.
2. In contrast to 1., full-time faculty size increased by only 22%, but their educational qualifications were up, with 11% holding a doctorate.
3. The faculty is young, with a median age of 40.
4. Enrollments in remedial mathematics increased by 81% and now account for 40% of all enrollments.
5. Hand calculators have gained widespread acceptance.

These were compared and contrasted with general trends in two-year colleges, with special attention to the occupational-technical areas.

The "Why Don't You Just...?" Barrier in Discrete Algorithms, by Professor H. S. Wilf, University of Pennsylvania.

The main theme of Professor Wilf's talk was the fact that when dealing with algorithms, it is important not only to present a good algorithm but also to discuss all of the other, simpler algorithms which come readily to mind, and to explain why they are inferior to the one presented. This idea was illustrated with algorithms from the theory of finite sets, partitions, graphs, etc.

Problems in Products, by Professor Mary Ellen Rudin, University of Wisconsin.

Which topological properties of factors are preserved in an ordinary Cartesian product of topological spaces? To what extent are the properties of finite subproducts preserved? If all factors are intervals what can we say about compact subsets of the product? A Stone-Čech compactification is such a product; many problems in this area are set theoretic. But it can be proved without set theoretic assumptions that there are remote points in $\beta\mathbf{R}$-$\mathbf{R}$.

Choosing a Wife, by Professor Leonard Gillman, The University of Texas, Austin. This talk has two purposes: to present some interesting mathematical reasoning, and to make a mathematical contribution in support of equal status for women. The problem is to determine a socially acceptable assignment of couples when each person in a community of n men and n women have ranked all those of the opposite sex in order of preference as a spouse. The methods are combinatorial. The material is based on the paper by David Gale and L. S. Shapley, *College Admissions and the Stability of Marriage*, this MONTHLY, January 1962, with embellishments by Donald Knuth (*Mariages Stables*, Les Presses de l'Université de Montréal, 1976) and the speaker.

The Optimal Distribution of Searching Effort, by Dr. B. O. Koopman, Adrain Professor Emeritus, Columbia University, and Arthur D. Little, Incorporated.

The talk was addressed to a class of methods for arriving at the optimum search for an object when only the probabilities of its various possible positions are known, and when the searching effort is of local effect, is subdivisible and additive. After analysing the physical pre-conditions for the mathematical formulation, analytical solutions are given in terms of the logarithm of the given probabilities, and their properties are studied. The treatment is elementary, but introduces such concepts as convexity, and inequality constraints. This theory, first developed by the author in a military context in 1946, has had wide civilian application.

SPECIAL SESSIONS OF THE ASSOCIATION

Film showings were held in the Chase Club Room of the Chase-Park Plaza Hotel on Friday and Sunday at 7:00 P.M. The following films were shown on Friday:

7:00–7:14 P.M.	SYMMETRIES OF THE CUBE, a film of the College Geometry Project
7:20–7:47 P.M.	MAURITS ESCHER: PAINTER OF FANTASIES
7:50–8:03 P.M.	MATHEMATICS ON THE HONEYCOMB
8:10–8:30 P.M.	TURNING A SPHERE INSIDE OUT, a film of the Topology Films Project
8:35–9:00 P.M.	THE GAUSS-BONNET THEOREM
	Films produced by T. F. Banchoff and Charles Strauss
9:05–9:10 P.M.	THE HYPERCUBE: PROJECTIONS AND SLICING (b&w)
9:15–9:20 P.M.	COMPLEX FUNCTIONS GRAPHED IN 4-SPACE
9:25–9:30 P.M.	THE VERONESE SURFACE GRAPHED IN 4-SPACE

The following films were shown on Sunday:

7:00–7:10 P.M. CAROMS, a film of the College Geometry Project
7:15–8:01 P.M. MATCHING THEORY — THE MARRIAGE THEOREM, Part I with Gian-Carlo Rota (b&w)
8:05–8:17 P.M. I MAXIMIZE
8:20–8:24 P.M. THE SEVEN BRIDGES OF KÖNIGSBERG, a film produced by Bruce and Katherine S. Cornwell
8:30–8:47 P.M. INFINITY
8:50–9:16 P.M. SPACE FILLING CURVES, a film of the Topology Films Project

MEETING OF THE BOARD OF GOVERNORS

The Board of Governors met on Friday, January 28, 1977, at 9:00 A.M. in the Empire Room of the Chase-Park Plaza Hotel with 40 members present. Among the items of business transacted were the following:

The Board elected H. E. Zink, Lane Community College, as Second Vice-President for the term 1977–79.

The Board elected T. E. Hull, University of Toronto, and Majorie L. Stein, U.S. Postal Service as Governors-at-Large for the term 1977–80. The Board had previously designated non-academic and Canadian mathematicians as two constituencies from which these Governors-at-Large were to be elected.

The Board approved the Nominating Committee for 1977 whose members are R. P. Boas, Chairman; D. J. Albers, and Lida K. Barrett.

The Board voted to approve the recommendation of the Committee on Earle Raymond Hedrick Lectures that Professor J. B. Keller, Courant Institute of Mathematical Sciences, be invited to deliver the twenty-fifth set of Earle Raymond Hedrick Lectures at the Association's meeting at the University of Washington, August 14–16, 1977. The Board elected Professor Martha Evens, Illinois Institute of Technology, as an Associate Editor of the AMERICAN MATHEMATICAL MONTHLY for the period extending through December 31, 1981.

Professor D. I. Schneider reported that the Committee on Educational Media has compiled a preliminary edition of "An Annotated Bibliography of Films and Videotapes for College Mathematics." A description, the name and address of the distributor, and review references are given for each of the films and video tapes listed.

Professor E. F. Beckenbach, Chairman of the Committee on Publications, reported that the following volumes are either in press or nearing completion: *Introduction to Field Theory* by C. E. Hadlock, *Studies in Combinatorics* by Gian-Carlo Rota, *Mathematical Methods in Science* by George Polya.

Professor D. B. Small, Chairman of the Committee on Secondary School Lectures, reported that EXXON has made a grant of $9350 to the Association to conduct a lectureship program designed to interest more Black students in mathematics. This program is to be conducted in the junior high schools of Atlanta, Washington, D.C., and Houston during the 1977–78 academic year. Officers of the National Association of Mathematicians are being consulted for assistance in conducting this program.

The Board voted to approve revised By-Laws of the New Jersey and New York Sections.

Professor W. F. Lucas reported that the volume, *Case Studies in Applied Mathematics*, had been mailed to chairmen of mathematics departments. Plans are also being made to publish a source book on mathematical applications in the social sciences.

The preliminary plans for the activities associated with the International Mathematical Olympiad were reported by the Secretary. These include an Awards Ceremony in Washington, D.C. sponsored by a grant of $7500 from IBM, a training session at the U.S. Military Academy, sponsored by a grant of $1800 from the Army Research Office, and travel to the 1977 International Mathematical Olympiad, sponsored by a grant of approximately $9000 from the Army Research Office.

The Board of Governors voted $4000 for support of a Congressional Science Fellow.

The following gifts to the Association were announced: a bequest of $47,462 from the estate of Professor Carl B. Allendoerfer, a gift of $2538 from Dorothy Allendoerfer, a gift of $50,000 from Carol Ryrie Brink, a gift of $1000 from Professor Harry M. Gehman, and anonymous gifts of $1000 and $500. The bequest from Professor Allendoerfer is to be used for support of the Association's publications and visiting lecturer programs.

The Board of Governors approved the operating budget for 1977 and discussed preliminary budgets for 1978 and 1979. The balanced budget for 1978 is based upon a dues increase which was approved by the Board and will be announced to the members in the 1977 summer dues notices.

The Board of Governors approved a new schedule of benefits for institutional members of the Association, effective beginning in 1978. There are to be two types of academic institutional members, *Academic Members* and *University Members.*

The schedule of benefits for an *Academic Member* is the following:

1. One subscription to each of the AMERICAN MATHEMATICAL MONTHLY, MATHEMATICS MAGAZINE, and the TWO-YEAR COLLEGE MATHEMATICS JOURNAL
2. Two copies of the COMBINED MEMBERSHIP LIST
3. One copy of the CUPM COMPENDIUM

4. The privilege of purchasing texts to be used for student awards at MAA members' prices
5. One institutional member (student or faculty)
6. Additional student membership for $12 per membership.

The schedule of benefits for *University Members* includes each of 1–6 as well as:

7. Five institutional student memberships

The 1978 dues for an *Academic Member* are $75 and for a *University Member* are $125.

The benefits for an *Academic Member* have value in excess of $100, and those of a *University Member* have value in excess of $175.

The Board of Governors approved Brown University and August 8–10, 1978, as the site and date of the Fifty-Eighth Summer Meeting of the Association. The complete schedule of meetings now approved follows:

University of Washington	August 14–16, 1977
Atlanta, Georgia	January 6–8, 1978
Brown University	August 8–10, 1978
Milwaukee, Wisconsin	January 13–15, 1979
Virginia Polytechnic Institute & State University	August 21–23, 1979
San Antonio, Texas	January 5–7, 1980
San Francisco, California	January 10–12, 1981

The Executive Director reported that the number of individual members in good standing as of December 9, 1976, was 18,201 compared with 18,648 on December 15, 1975. The number of Academic members was 416, compared to 401 one year earlier. The Executive Director also reported the following circulation figures for the three journals published by the Association: AMERICAN MATHEMATICAL MONTHLY, 19,068; MATHEMATICS MAGAZINE 7859; TWO-YEAR COLLEGE MATHEMATICS JOURNAL 4927.

The Board of Governors approved a statement on the preparation expected of students planning to take collegiate level courses and eleven recommendations designed to reverse the decline. This material was prepared by the Committee on the Reported Decline in the Preparation of Students for Collegiate Mathematics. The same statement and recommendations have been submitted to NCTM for approval.

The Board of Governors approved the establishment of an MAA award, the Certificate of Merit. The guidelines for these Certificates follow:

1. The MAA shall establish awards of Certificates of Merit to be presented to individuals who have made exemplary contributions to the MAA or the mathematical community in general. These contributions may be of any type whatsoever, but must have been in furtherance of the objectives of the MAA or the best interests of the mathematical community in general.

2. The recipients would typically be those who have made contributions of the type for which honorary life memberships have been awarded in the past. The Executive and Finance Committees may decide to award such certificates whenever they believe them to be appropriate, but normally no more than 2 Certificates of Merit shall be presented during any one year. The award is not intended to be one where the Executive and Finance Committees would actively look for recipients; rather it is intended to be a mechanism to have available when the need to make an award arises spontaneously.

3. The decision of the Executive and Finance Committees shall be final and not require confirmation by the Board of Governors.

4. The recipients will be presented with an appropriately inscribed certificate at a business meeting of the Association.

ANNUAL BUSINESS MEETING OF THE ASSOCIATION

The Annual Business Meeting of the Association was held on Sunday, January 30, 1977, in the Khorassan Ballroom of the Chase-Park Plaza Hotel, with President Pollak presiding.

The Association's Sixteenth Award for Distinguished Service was made to Professor Victor L. Klee, Jr. of the University of Washington. The citation (which appears on pages 81–82 of the February issue of the MONTHLY) was prepared and read by Dr. Truman A. Botts. Professor Klee, in accepting the Award, expressed his appreciation to the Association for the Award, to the mathematical community for the opportunity to serve, and to his wife for her patience and moral support during that period.

A specially-bound copy of the citation was presented to Mrs. Klee in appreciation of her support of Professor Klee's many activities.

The Chauvenet Prize for 1977 was presented to Professor W. Gilbert Strang for his paper, "Piecewise Polynomials and the Finite Element Method," which appeared in the BULLETIN OF THE AMERICAN MATHEMATICAL SOCIETY, 79 (1973), 1128–37. Further details concerning this Prize and its recipient will appear in the June-July issue of the MONTHLY and in the June issue of the TYCMJ.

Professor Strang, in accepting the Prize, commented as follows:

"I am extremely grateful for the award of the Chauvenet Prize. I have always felt that we each could look for a way to write about mathematics that is lively and clear, but, of course, I never expected such an honor."

Professor Strang also expressed his happiest thanks to his teachers and friends (and in particular to Peter Lax and Erwin Canham) for their encouragement over many years.

President Pollak next announced that Professor Henry M. Cox of the University of Nebraska was to be presented the Association's first Certificate of Merit. He called on Professor W. E. Mientka who read the following citation:

"Throughout the mathematical community particular individuals are recognized for their contributions to Instructional Research. One of these is Professor Henry Miot Cox, to whom the MAA today presents a Certificate of Merit.

Henry Miot Cox was born in Toccoa, Georgia and he received a Bachelor of Science degree from Emory University and a Master of Arts degree from Duke University. Further graduate work in mathematics was completed at Princeton University. From 1930–1939 Henry held teaching and administrative positions with the State of Georgia University System; he taught at the Georgia School of Technology and the University of Georgia and served as the Assistant Examiner for the University System. In 1939 he was employed by the University of Nebraska and was appointed Director of the Bureau of Instructional Research and Examination Services in 1940. He served in this position until his retirement in 1973. The Bureau served as a research service agency of the University, working cooperatively with administrative offices and with instructional departments. In addition to responsibilities and services within the University, the Bureau extended assistance in the interpretation of examinations to Nebraska high schools and colleges, as well as to faculty members and to graduate students engaged in individual research.

During Henry's tenure as Director of the Bureau of Instructional Research and Examination Services he expanded the set of classification examinations which included examinations in the natural sciences, social studies, vocabulary (linguistic ability), quantitative ability, reading speed and reading comprehension. The results of these examinations were used to guide entering students to appropriate beginning academic courses and they were subsequently used to compare the academic performance of the students in the same subject areas upon graduation. The examinations were also used to determine the University's Regents Scholarship winners. Henry's research results have been published in several Journals including the *Journal of the American Association of Collegiate Registrars, The Journal of Education Psychology* and the *American Mathematical Monthly.*

Henry's expertise has often been solicited by other organizations. He has served as a member of the Mathematics Department at the University of Nebraska, Nebraska Section of the Mathematical Association of America (he served as Chairman of the Section and from 1957 to the present as the Secretary/Treasurer) and various civic organizations including the Nebraska Bicentennial Commission and the Sons of the American Revolution. In addition Henry has been an active layman in the Methodist church and a student of genealogy since the age of seventeen.

In 1957 a group of Nebraska High School and University Mathematics Teachers formed an *ad hoc* committee which met to discuss the possibility of participating in the Annual High School Mathematics Examination. Henry was a member of this committee and his expertise with examinations was very much appreciated by the committee during its deliberations.

On August 15, 1970 the Association appointed Henry to succeed Professor W. H. Fagerstrom as the Executive Director of the Annual High School Mathematics Examination. During the six years he served in this position Henry established effective administrative procedures and gained the respect and appreciation of the sixty-one Regional Contest Chairmen associated with the Examination. He established excellent rapport with the officers of the MAA and the Examination sponsors. Henry devoted many hours to the examination process and personally visited many of the Regional Contest Chairmen in order to provide direct assistance. During Henry's tenure as the Executive Director the number of students who participated in the Examination increased from 329,000 to 365,000.

Henry is known as a warm and friendly person and has gained a reputation for his meticulous and precise mode of operation. He began his teaching experience with a position at Emory Academy and his teaching assignment included Mathematics, Physics, General Science and English. In addition he was responsible for the periodic supervision of the gymnasium and was required to be available to counsel students twenty-four hours a day, seven days a week. This early indication of Henry's versatility was a prelude to the future.

The Mathematical Association of America extends to Professor Henry Miot Cox a Certificate of Merit in recognition of and with thanks for his many years of distinguished service to Instructional Research and to the Nebraska Section of the Mathematical Association of America and to the Annual High School Mathematics Examination."

Professor Cox was unable to attend the meeting, but sent the following letter which was read by the Secretary:

"I accept the MAA Certificate of Merit with thanks to the officers and members of the Mathematical Association of America who have assisted in so many ways the work of the Annual High School Mathematics

Examination. I accept as a representative of the hundreds of college teachers, the thousands of high school teachers, and the hundreds of thousands of high school students who each year take part in the program. Annually we have met the goal which we have set, namely to identify and to recognize young people, boys and girls, who have mathematical ability and mathematical interests. These young people represent our hope and our promise for the future."

President Pollak thanked Professor Betty J. Hinman, whose term as Second Vice-President expired at this meeting, for her dedicated service to the Association during the previous two years.

The Secretary presented his report announcing first some of the actions taken by the Board of Governors, the membership figures, and the budgetary situation of the Association. The schedule of benefits for Academic and University Members beginning in 1978 was given and the increase in dues for individual members beginning in 1978 was detailed.

The Secretary called attention to the very fine program at this meeting. He indicated that Professor G. E. Weiss, Chairman of the Program Committee, was deserving of the thanks of all of the members of the Association for having so carefully organized this program. The Secretary especially thanked Professor Leonard Gillman, Treasurer of the Association, for having agreed to give an hour address in place of Professor S. M. Ulam, who was forced to cancel his address.

The Secretary also thanked the members of the Committee on Arrangements for having arranged the meeting. In particular, the Secretary thanked Professor J. N. Siegel of the University of Missouri at St. Louis for having chaired the Committee on Arrangements. The Secretary noted that there had been several complicated arrangements for this meeting which he thanked the Committee on Arrangements for having handled. In particular, the Secretary said that the AMS-MAA concert and the display of combinatorial games arranged for the AMS had been particularly difficult. At the request of AMS President R. H. Bing, the Secretary thanked the Arrangements Committee on behalf of both the AMS and the MAA members.

MEETINGS OF OTHER ORGANIZATIONS

On Wednesday evening the AMS-SIAM Committee on Mathematics in the Life Sciences presented a symposium on Mathematical Ecology. Speakers were S. A. Levin, R. M. May and George Oster.

The AMS held sessions from Thursday, January 27, through Sunday, January 30. The following AMS addresses were given in Khorassan Ballroom (arranged in chronological order of presentation):

THURSDAY

9:30 A.M. *The Integrability Problem for Lie Equations*, Professor H. L. Goldschmidt, Princeton University.

11:00 A.M. *Inequalities in Fourier Analysis*, Professor William Beckner, University of Chicago.

1:00 P.M. AMS Colloquium Lecture: *Differential Topology of Higher Dimensional Manifolds*, Professor William Browder, Princeton University.

2:15 P.M. *Variational Inequalities and Related Free Boundary Value Problems*, Professor D. S. Kinderlehrer, University of Minnesota.

3:30 P.M. *G-Smoothing Theory*, Professor R. K. Lashof, University of Chicago.

4:45 P.M. *Quasi Monte Carlo Methods and Pseudo-Random Numbers — Some Applications of Number Theory*, Professor H. G. Niederreiter, University of Illinois.

8:30 P.M. Fiftieth Joshiah Willard Gibbs Lecture: *Rays, Waves, and Asymptotics*, Professor J. B. Keller, New York University.

FRIDAY

8:30 A.M. *Bases, Biorthogonal Systems, and Various Approximation Structures in Banach Spaces*, Professor Aleksander Pelczynski, Polish Academy of Science.

9:45 A.M. *Hyperelliptic Riemann Surfaces with Infinitely Many Branch Points and the Related Theta Function*, Professor H. P. McKean, Jr., Courant Institute of Mathematical Sciences.

11:00 A.M. *Mathematical Physics of Phase Transitions and Symmetry Breaking: New Rigorous Results*, Professor Jurg Frohlich, Princeton University.

1:00 P.M. AMS Colloquium Lecture II: Professor William Browder.

2:15 P.M. AMS Retiring Presidential Address: *Quasiconformal Mappings with Applications to Differential Equations, Function Theory, and Topology: A Report for Nonspecialists*, Professor Lipman Bers, Columbia University.

3:30 P.M. Award of the Frank Nelson Cole Prize in Number Theory.

4:00 P.M. AMS Business Meeting.

5:30 P.M. Open meeting on the job market sponsored by AMS Committee on Employment and Educational Policy (CEEP).

8:00 P.M. CEEP Panel: *Graduate Education*, Professors P. T. Bateman, moderator, Lipman Bers, R. H. Bing, and G. D. Mostow.

SATURDAY
1:00 P.M. AMS Colloquium Lecture III: Professor William Browder.
2:15 P.M. *Beyond First-Order Logic*, Professor K. J. Barwise, University of Wisconsin.
3:30 P.M. *Monodromy Groups and Poincaré Series*, Professor D. A. Hejhal, Columbia University.

SUNDAY
1:00 P.M. AMS Colloquium Lecture IV: Professor William Browder.
2:15 P.M *An Algebraic Approach to the Topological Degree of a C^∞ Map*, Professor David Eisenbud, Brandeis University.

The Association for Women in Mathematics (AWM) sponsored a panel discussion on "More History of Women in Mathematics" at noon on Sunday in the Chase Club Room of the Chase-Park Plaza Hotel. Professor Leonore Blum served as moderator. There was an open meeting of the AWM Executive Committee Saturday at 4:30 P.M., and an AWM discussion session at 5:30 P.M. on the same day.

The Conference Board of the Mathematical Sciences (CBMS) sponsored a panel discussion on "Environmental Health and Mathematics" at 2:30 P.M. on Saturday in Moore Auditorium of the Washington University Medical School. Dr. D. L. Thomsen, Jr., of the SIAM Institute for Mathematics and Society, served as moderator. Panelists included Professor K. B. Bischoff, University of Delaware; Norman Breslow, University of Washington; P. S. Gartside, University of Cincinnati. The CBMS Council met at 2:15 P.M. on Sunday in the Carriage Room of the Bel Air West.

The Mathematicians Action Group (MAG) held an open meeting of its Steering Committee at 9:00 A.M. on Wednesday in the Venetian Room of the Chase-Park Plaza. The MAG Business Meeting was at 3:00 P.M. on Thursday. MAG also sponsored a panel discussion at 8:30 A.M. on Sunday in the Empire Room of the Chase-Park Plaza.

At 7:00 P.M. on Thursday in the Stockholm Room of the Chase-Park Plaza, Professor Raymond Johnson delivered an hour address to the National Association of Mathematicians (NAM). On Friday at 7:15 P.M. in the Phillip Room of the Bel Air West, NAM sponsored a panel discussion, "Some Different Approaches to the Teaching of Mathematics-Success Oriented Learning Principles and Techniques." Panelists were Professors E. M. Carroll, J. E. Hall, Frank Hawkins, Grady Nelson, and Louis Richards, moderator.

The Association for Symbolic Logic held sessions on Thursday and Friday. Invited ASL addresses will be delivered by Professor R. L. Vaught and by Professor C. G. Jockusch, Jr. on Thursday at 10:45 A.M. and 1:30 P.M., respectively, and Fred Galvin at 1:30 P.M. on Friday.

ARRANGEMENTS, ENTERTAINMENT, AND RECREATION

The Committee on Arrangements consisted of J. N. Siegel, Chairman; P. T. Bateman, C. M. Bruns, W. C. Connett, R. C. Freiwald, Kenneth Kunen, J. B. Riles, D. P. Roselle, Nadine L. Verderber, G. L. Walker.

Registration headquarters were located in the Chase Club Lounge of the Chase-Park Plaza Hotel on Wednesday and in Exhibit Hall A Thursday through Monday. The Employment Register was held in Exhibit Hall C of the Chase-Park Plaza on Friday, Saturday and Sunday with interviews scheduled from 9:00 A.M. to 5:30 P.M. on Saturday and Sunday. Books and educational media exhibits were on exhibit in Exhibit Hall A of the Chase-Park Plaza from 1:00–5:00 P.M. on Thursday, 9:00 A.M.–5:00 P.M. on Friday and Saturday, and 9:00 A.M.–Noon on Sunday.

At 9:00 P.M. on Friday, January 29, there was a cocktail party featuring Dixieland jazz in the Starlight Room at the Chase-Park Plaza Hotel.

DAVID P. ROSELLE, *Secretary*

OCTOBER MEETING OF THE OHIO SECTION

The Ohio Section of the MAA held its Fall meeting at Marshall University, Huntington, West Virginia, October 22 and 23, 1976. Approximately one hundred and eighty people were in attendance. Section Chairman J. A. Murtha presided; M. D. Wetzel was the Program Chairperson. The theme for the meeting centered upon "Mathematical Modeling".

Invited addresses included: "Graphs, Garbage, and a Pollution Solution: Graph Theory Applied to Environmental Problems", by F. S. Roberts, Rutgers University; and "Discrete Mathematical Models for Some Problems Arising in Biology and Medicine", by Maynard Thompson, Indiana University. A highlight of the meeting was the "Swap Session on Mathematical Modeling", moderated by Professors Roberts and Thompson, and by R. L. Wilson, Ohio Sectional Governor.

The following contributed papers were presented:

Banzhaf Values of the 1976 Presidential Election Game, by E. Bolger, Miami University.

A Model of Downstream Carriage of Pesticides in Natural Streams, by P. Gingo, University of Akron.

A Comparison of Several Non-Parametric Measures of Location, by P. Greenough, Marshall University.

The Geological Environment Affecting a Model of Leachate Dispersion from a Landfill, by J. Jackson and P. Gingo, University of Akron.

Use of Computers in Modeling, by Z. Karian, Denison University.

The Limiting Behavior for Several Interacting Populations, by A. Leung, University of Cincinnati.

Women in Mathematics — Changing Times, Changing Attitudes, by M. Levine and R. Rolwing, University of Cincinnati.

On Identifying Models of the Human, by D. Norris, Ohio University.

Numerical Solution of Optimal Vaccination Schemes, by V. Norton, Bowling Green State University.

Modeling in Baseball, by M. Pankin, Marshall University.

The Problem with Secretaries, by C. Pinzka, University of Cincinnati.

Teaching Applications to Undergraduates — More New Modular Materials, by P. Tuchinsky, Ohio Wesleyan University.

The agenda also included a brief Business Meeting of the Section, a meeting of the Executive Committee, and meetings of ad hoc committees: Committee on Cooperation among Colleges and Universities, Committee on Curriculum, and Committee on Teacher Training and Certification.

GUS MAVRIGIAN, *Secretary-Treasurer*

NOVEMBER MEETING OF THE PHILADELPHIA SECTION

The fifty-first annual meeting of the Philadelphia Section of the MAA was held on November 20, 1976, at Montgomery County Community College, Blue Bell, Pennsylvania. Section Chairman, Professor Eugene Klotz, and Section Vice-chairman, Professor Doris Schattschneider, presided. A total of 209 persons attended, including 149 members of the Association.

The following papers were presented:

Proving the Four Color Theorem, by John Koch, Wilkes College.

The Sound of Mathematics, by Samuel Plotkin and John Clark, Montgomery County Community College.

Mathematical Aspects of Periodic Catatonic Schizophrenia, by Jane Cronin, Rutgers University.

Catastrophe Theory and Its Application, by Nelson Max, Case Western Reserve University.

A special session for students heard the following papers:

Reflections on a Counting Problem: An Application of Symmetry Groups, by David Bayer, Swarthmore College.

Computerizing Readability Formulas, by Paul Haas, Moravian College.

The top performer from the Section in the 1975 Putnam Examination was Dennis DeTurck of Drexel University. He was awarded a membership for one year in the MAA.

Section Officers elected for 1977 are: E. A. Klotz, Swarthmore College, Chairman; D. W. Schattschneider, Moravian College, Vice-Chairman; W. Baxter, University of Delaware, Secretary-Treasurer. Newly elected members of the Executive Committee are: L. K. Arnold, Daniel H. Wagner Associates; J. M. Rubillo, Bucks County Community College.

P. E. BEDIENT, *Secretary-Treasurer*

CALENDAR OF FUTURE MEETINGS

Fifty-seventh Summer Meeting, University of Washington, August 14–16, 1977.

Sixty-first Annual Meeting, Atlanta, Georgia, January 6–8, 1978.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN, last weekend in April or first weekend in May. Deadline for papers 6 wks. bef. mtg.

FLORIDA, early March. Deadline for paper titles 2 wks. bef. mtg.

ILLINOIS, Chicago Loop College, Chicago, May 6–7, 1977.

INDIANA

INTERMOUNTAIN

IOWA, third weekend in April. Deadline for papers February 1.

KANSAS, March or April. Deadline for papers January 1.

KENTUCKY, early April. Deadline for papers 6 wks. bef. mtg.

LOUISIANA-MISSISSIPPI, Buena Vista Hotel-Motel, Biloxi, Mississippi, February 17–18, 1978.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA, Saturday before Thanksgiving and last Saturday in April.

METROPOLITAN NEW YORK, Spring. Deadline for papers 2 wks. bef. mtg.

MICHIGAN, Eastern Michigan University, Ypsilanti, May 6–7, 1977.

MISSOURI, late March/early April. Deadline for papers January 31.

NEBRASKA, April.

NEW JERSEY, early November and early May.

NORTH CENTRAL, end of October and April. Deadline for papers October 1 and April 1.

NORTHEASTERN, Middlebury College, Middlebury, Vermont, June 18–19, 1977.

NORTHERN CALIFORNIA, College of Notre Dame, Belmont, February 1978.

OHIO

OKLAHOMA-ARKANSAS, (approx.) Friday and Saturday of first weekend in April. Deadline for papers 3 wks. bef. mtg.

PACIFIC NORTHWEST, second Saturday in June. Deadline for papers 6 wks. bef. mtg.

PHILADELPHIA, Moravian College, Bethlehem, November 19, 1977.

ROCKY MOUNTAIN, last weekend in April or first in May. Deadline for papers 8 wks. bef. mtg.

SEAWAY, State University College at Buffalo, May 6–7, 1977.

SOUTHEASTERN

SOUTHERN CALIFORNIA, first or second Saturday in March.

SOUTHWESTERN, usually in April. Deadline for papers 2 wks. bef. mtg.

TEXAS, Friday and Saturday in early April. Deadline for papers March 1.

WISCONSIN, Friday and Saturday between mid-April and first week in May. Deadline for papers 6 wks. bef. mtg.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Washington, February 12–17, 1978.

AMERICAN MATHEMATICAL ASSOCIATION OF TWO YEAR COLLEGES

AMERICAN MATHEMATICAL SOCIETY, University of Washington, August 15–18, 1977.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION, University of North Dakota, Grand Forks, June 13–16, 1977.

ASSOCIATION FOR COMPUTING MACHINERY, Olympic Hotel, Seattle, Washington, October 17–19, 1977.

ASSOCIATION FOR SYMBOLIC LOGIC, Wrocław, Poland, August 1–2, 1977.

ASSOCIATION FOR WOMEN IN MATHEMATICS

CANADIAN SOCIETY FOR HISTORY AND PHILOSOPHY OF MATHEMATICS/SOCIÉTÉ CANADIENNE D'HISTOIRE ET DE PHILOSOPHIE DES MATHÉMATIQUES, Hamilton, Ontario, June 2, 1977.

FIBONACCI ASSOCIATION

INSTITUTE OF MATHEMATICAL STATISTICS, Seattle, Washington, August 14–18, 1977.

MU ALPHA THETA, Loras College, Dubuque, Iowa, August 7–10, 1977.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, San Diego, California, April 12–15, 1978.

OPERATIONS RESEARCH SOCIETY OF AMERICA, San Francisco Hilton, May 9–11, 1977.

PI MU EPSILON, University of Washington, August 14–16, 1977.

SCHOOL SCIENCE AND MATHEMATICS ASSOCIATION, William Penn Hotel, Pittsburgh, Pennsylvania, November 10–12, 1977.

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, Philadelphia-Sheraton Hotel, Philadelphia, June 13–15, 1977 (25th Anniversary Meeting).

THE CARUS MATHEMATICAL MONOGRAPHS

The Monographs are a series of expository books intended to make topics in pure and applied mathematics accessible to teachers and students of mathematics and also to non-specialists and scientific workers in other fields.

These numbers are currently available:

1. *Calculus of Variations,* by G. A. Bliss.
2. *Analytic Functions of a Complex Variable,* by D. R. Curtiss.
3. *Mathematical Statistics,* by H. L. Rietz.
4. *Projective Geometry,* by J. W. Young.
6. *Fourier Series and Orthogonal Polynomials,* by Dunham Jackson.
7. *Vectors and Matrices,* by C. C. MacDuffee.
8. *Rings and Ideals,* by N. H. McCoy.
9. *The Theory of Algebraic Numbers* (Second edition), by Harry Pollard and Harold G. Diamond.
10. *The Arithmetic Theory of Quadratic Forms,* by B. W. Jones.
11. *Irrational Numbers,* by Ivan Niven.
12. *Statistical Independence in Probability, Analysis and Number Theory,* by Mark Kac.
13. *A Primer of Real Functions* (Second edition), by Ralph P. Boas, Jr.
14. *Combinatorial Mathematics,* by H. J. Ryser.
15. *Noncommutative Rings,* by I. N. Herstein.
16. *Dedekind Sums,* by Hans Rademacher and Emil Grosswald.
17. *The Schwarz Function and its Applications,* by Philip J. Davis.
18. *Celestial Mechanics,* by Harry Pollard.

One copy of each Carus Monograph may be purchased by individual members of the Association for $6.50 each; additional copies and copies for nonmembers are priced at $11.00 each. (Orders for under $10.00 must be accompanied by payment. Prepaid orders will be delivered postage and handling free.)

Orders should be sent to:

MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.
Washington, D.C. 20036

New from Wiley-Interscience

APPLIED NONSTANDARD ANALYSIS

Martin Davis
This applications-oriented work is intended to make the methods of nonstandard analysis available to readers with no previous knowledge of logic
181 pp (0 471 19897-8) **1977** $16 95

STATISTICAL METHODS FOR DIGITAL COMPUTERS
Vol. 3 of Mathematical Methods for Digital Computers

Edited by **Kurt Enslein, Anthony Ralston & Herbert S. Wilf**
Nineteen contributions on major multivariate statistical techniques used by statisticians Gives background as well as computer aspects of common statistical computation techniques
432 pp (0 471 70690-6) **1977** $24 95

ORDINARY DIFFERENTIAL EQUATIONS IN THE COMPLEX DOMAIN

Einar Hille
Gives exclusive coverage on solutions in the complex domain Incorporates much material not available in other treatises Gives full coverage to existence theorems, representation of solutions by series integrals application of the modern value distribution theory and more
484 pp (0 471 39964-7) **1976** $27 95

APPLIED AND COMPUTATIONAL COMPLEX ANALYSIS, Vol. 2

Peter Henrici
A self-contained presentation of major areas of complex analysis Topics covered include special functions, integral transforms, asymptotic expansions, and continued fractions
Approx 672 pp (0 471 01525-3) **1976** $25 00

ALGEBRA, Vol. 2

P.M. Cohn
This text discusses topics usually found in an advanced undergraduate or postgraduate course in modern algebra and includes homological algebra, lattices and non-commutative Noetherian rings The material falls roughly into three parts basic theories, fields and rings
483 pp (0 471 01823-6) **1976** $19 50

MATHEMATICAL BIOECONOMICS
The Optimal Management of Renewable Resources

Colin W. Clark
A unified treatment of resource management techniques based on contemporary mathematical and economic methods Employs mathematical models such as those used in fisheries, forestry, and wildlife management, applying modern principles of cost-benefit analysis to determine optimal exploitation policies
368 pp (0 471 15856-9) **1976** $21 95

MODERN ALGEBRA WITH APPLICATIONS

William J. Gilbert
Presents all group, ring, and field theory necessary for a first course in modern algebra Includes monoids and semigroups, Boolean algebras and switching circuits, and Polya-Burnside theory
348 pp (0 471 29891-3) **1976** $18 95

COMPUTATIONAL ANALYSIS WITH THE HP-25 POCKET CALCULATOR

Peter Henrici
Contains 35 high-level mathematical programs written for a specific programmable pocket calculator, the HP-25 These programs implement algorithms in number theory, equation solving, algebraic stability theory, calculus of power series and numerical integration, as well as algorithms for the evaluation of special higher transcendental functions such as the Gamma function, various Bessel functions, the error integral, and the Riemann zeta function Programs will adapt to run on any calculator of comparable capacity
Approx 160 pp (0 471 02938-6) **1977**
$9 95 (tent)

NONLINEAR OPERATORS AND DIFFERENTIAL EQUATIONS IN BANACH SPACES

Robert H. Martin, Jr.
Organizes and develops many basic and important techniques in Nonlinear Functional Analysis with emphasis on *nonlinear* methods Examines and applies new techniques including the measure of noncompactness
440 pp (0 471 57363-9) **1976** $27 50

MATHEMATICAL IDEAS, MODELING AND APPLICATIONS
Vol. 2 of Companion to Concrete Mathematics

Z.A. Melzak
Contents Further Topics in Geometry Analysis Topics in Combinatorics Number Theory and Algebra An Approach to Computing and Computability Topics in Applications Functional Equations and Mathematical Modeling Appendix Bibliography Index
413 pp (0 471 59341-9) **1976** $29 95

Available at your bookstore or write to Nat Bodian
Dept A8032-51

WILEY-INTERSCIENCE
a division of John Wiley & Sons, Inc
605 Third Avenue
New York, N Y 10016
In Canada 22 Worcester Road, Rexdale, Ontario
Prices subject to change without notice

092 A8032-51

THE AMERICAN

MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

AN OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 84 NUMBER 6

CODEN: AMMYAE

CONTENTS

(Continued on inside cover)

JUNE-JULY 1977

NOTICE TO AUTHORS

Research papers per se are unsuitable; see statement of policy (Vol. 81, p. 1).

Please follow the format in current issues of the MONTHLY. Manuscripts must be legibly typewritten or reproduced from typewritten copy, double spaced with wide margins and on one side of the paper. Three copies should be submitted to the appropriate editor and one kept by the author as protection against loss. The author's full address *must* appear at the end of the manuscript.

Backlog: Main Articles 17 months, Math. Notes 17 months, Research Problems 12 months, Classroom Notes 12 months, Math. Education 12 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to R. P. BOAS, Department of Mathematics, Northwestern University, Evanston, IL 60201; NOTES, etc.: to the corresponding Associate Editor; REPRINT PERMISSION: to LEONARD GILLMAN, Mathematical Association of America, University of Texas, Austin, Texas 78712 (see also the copyright notice below); ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N.Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N. W., Washington, D.C. 20036; BACK ISSUES: Contact P. and H. Bliss Co., Middletown, CT 06457.

Annual dues for members of the Association, including a subscription to the American Mathematical Monthly, are $18.00 for each of the first two years of membership and $22.00 thereafter. Student Membership is available with annual dues of $12.00. For nonmembers the subscription price is $25.00.

PUBLISHED BY THE ASSOCIATION at Washington, D.C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D.C., and additional mailing offices.

PRINTED IN THE UNITED STATES OF AMERICA

AWARD OF THE 1977 CHAUVENET PRIZE TO PROFESSOR GILBERT STRANG

The Board of Governors of the Mathematical Association of America voted to award the 1977 Chauvenet Prize to Professor Gilbert Strang for his paper, "Piecewise Polynomials and the Finite Element Method," Bulletin of the American Mathematical Society, 79 (1973) 1128–37.

A certificate and monetary award in the amount of five hundred dollars was presented to Professor Strang at the Business Meeting of the Association on January 30, 1977 in St. Louis, Missouri.

The Chauvenet Prize is awarded for a noteworthy paper of an expository or survey nature published in English, which comes within the range of profitable reading for members of the Association. The purpose of the Prize is to stimulate the writing of expository and survey articles. The 1977 Prize, awarded for a paper published in the three year period 1973–75, is the twenty-fifth award of the Chauvenet Prize since its institution by the Association in 1925. For the list of names of the previous winners, see this MONTHLY, 71 (1964), p. 589; 72 (1965), pp. 2–3; 74 (1967), p. 3; 75 (1968), pp. 3–4; 77 (1970), pp. 117–118; 78 (1971), pp. 112–113; 79 (1972), pp. 112–113; 80 (1973), pp. 120; 81 (1974), pp. 113–114; 82 (1975), pp. 108–109; and 83 (1976), pp. 84–85.

Professor Strang was born on November 27, 1934, in Chicago. He received the S.B. degree from M.I.T. in 1955, and was elected to a Rhodes Scholarship for postgraduate study in England. He was awarded First Class Honours at Balliol College, Oxford in 1957. After completing the Ph.D. degree at U.C.L.A. in 1959, under the supervision of Professor Peter Henrici, he returned to M.I.T. as a Moore Instructor and has remained a member of the M.I.T. faculty ever since. He was promoted to Professor in 1970, and became chairman of the Committee on Pure Mathematics in 1975.

Professor Strang was recently elected to the Council of the Society for Industrial and Applied Mathematics, after service on the AMS-SIAM Committee on Applied Mathematics. He was chosen for a NATO Fellowship in 1961 and for a Sloan Fellowship in 1966.

His mathematical interests have centered on partial differential equations and their discrete approximations. His earliest papers studied the stability of different methods for the Cauchy problem, and with Hermann Flaschka he investigated correctness in the presence of multiple characteristics. Professor Strang's more recent work has been concerned especially with the finite element method, which was the subject of a lecture to the American Mathematical Society and of the paper which has been awarded the Chauvenet Prize. With George Fix, he is the author of *An Analysis of the Finite Element Method* (Prentice-Hall, 1973). His teaching and his work with students have concentrated very strongly on applied linear algebra, and they led in 1976 to a new undergraduate textbook on *Linear Algebra and Its Applications* (Academic Press).

In his acceptance, Professor Strang expressed his gratitude to the Association for the honor of the 1977 Chauvenet Prize, and his happiest thanks to teachers and friends (and in particular to Peter Lax and Erwin Canham) for their encouragement over many years.

D.P. ROSELLE

ALL GAMES BRIGHT AND BEAUTIFUL

J. H. CONWAY

Our topic is the addition theory of partizan games. This means that although this paper is written after [3], it should naturally precede [3] on grounds of logic as well as euphony, since the number system of [3] was in fact suggested by the more general system described here. I thank Donald Knuth for suggesting that I write a survey paper with this title.

No proofs will be found in this paper. We hope that most readers will be interested enough to prove the more basic results for themselves, and rich enough to buy at least one of [1] and [4] if they find themselves stuck with the more difficult ones. The games described here are all treated more fully in [1] or [4], and in many cases their descriptions are taken almost verbatim from one of those two references.

To see how addition of games comes about, we consider two particular cases.

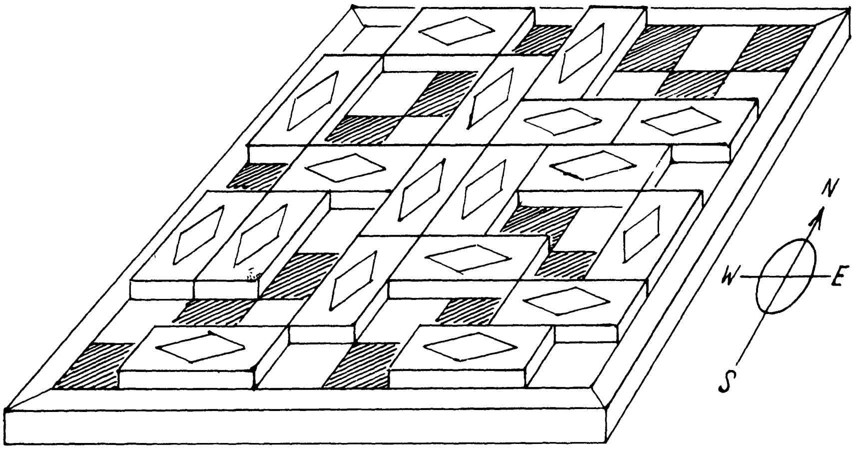


FIG. 1. A game of Domineering.

Domineering (proposed by Göran Andersson). This is played on a checkerboard, with a number of dominoes each of which can cover exactly two squares of the board. The player called Left, when it is his turn to move, must place a domino in North-South orientation so as to cover two currently empty squares, while Right, at each of his turns, places a domino oriented East-West, again so as to

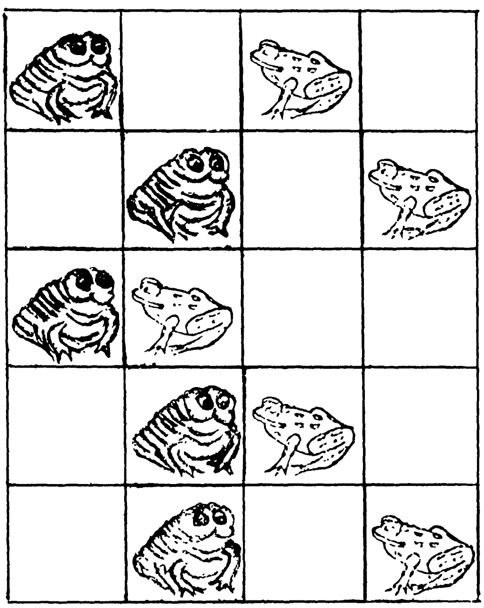


FIG. 2. A game of Toads and Frogs.

exactly cover two previously empty squares. If either player, when it is his turn to move, finds himself unable to place a domino in the required orientation, that player loses.

Toads and Frogs (proposed by Richard Guy). In Figure 2, Left has trained a number of Toads (*Bufo bufo*), and Right a number of Frogs (*Rana rana*) to play the following game. When it is Left's turn to move, he must *either* make one of his toads move just one place Eastward onto an empty square of the board, *or* persuade some toad to jump over a frog just to the East of it and land on the square just beyond that frog, which must be empty. Right moves his frogs in a similar way, but in the Westward direction. If a player whose turn it is to move cannot move any of his creatures in the prescribed way, then that player loses.

Partizan games in general. We now use these examples to illustrate our terminology. Both games are played according to the *normal play convention*, according to which a player loses *if and only if* he is unable to move when it is his turn to do so. And both satisfy the *finishing condition*, that there can be no infinite sequence of legal moves, whether made alternately by the two players or not. From now on, we shall understand that these two conditions apply to every game we consider. Note that the normal play convention enables us to avoid defining the winner in each individual case—he is simply that player who makes the last move of the game, and of course the finishing condition ensures that there will always be a last move. It should soon become obvious why we do not restrict the finishing condition to alternating sequences only. Because we do *not* impose the frequently added restriction that exactly the same moves are available to each player, we shall refer to our games as *partizan games* (and deliberately use the less common spelling of this word).

Sums of partizan games. The most interesting feature of our two examples is the way their typical positions break up into sums of rather smaller positions. Thus the available space in the Domineering position of Figure 1 is composed of separate regions of the shapes shown in Figure 3. When it is some

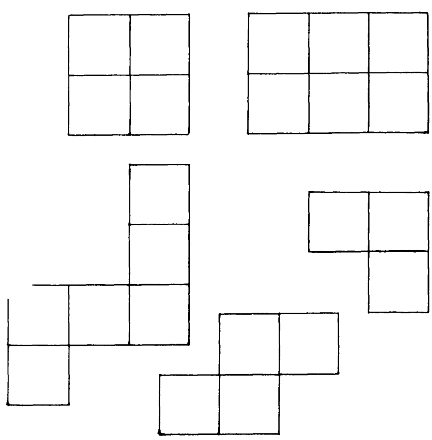


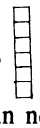
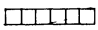
FIG. 3. Available regions in the Domineering position of Figure 1.

player's turn to move, he is forced to choose just one of these regions, and make a move legal for him in that region, and moves made in one region do not interfere with those made in another. Again, in a Toads and Frogs game, a player must choose just one of the East-West lanes into which the board is divided, and make a move legal for him in that lane, and moves made in one lane will not interfere with those made in another.

More generally, we can play the (*disjunctive*) *sum* of any finite collection of partizan games, the individual games being called the *components* of the sum. Each player, when it is his turn to move,

must choose just one of the components and make a move legal for him in that component. If he finds that there is no component in which he has a legal move, then of course he loses, by the normal play convention. The moves of the sum that affect any particular component need not of course be made alternately by the two players, but the strong finishing condition that we have adopted still ensures that the sum will necessarily end. In fact the finishing condition on the components is enough to ensure the finishing condition on their sum, which is therefore another partizan game.

Evaluating positions. It turns out that it is possible to assign *values* to the positions of partizan games in such a way that the value assigned to a sum of games is just the sum of the values of the components. In many cases the values are ordinary numbers, which are added according to the rules we learned at school, but in many more cases they are not. The theory of partizan games is concerned with the rules for adding and comparing the many other weird and wonderful values that arise.

Integer values. The positions  ,  ,  , in Domineering are easy to evaluate, since Right, who must place his Dominoes East-West, can never move, and Left has at most 1, 2, 3 moves in the respective cases. Since we shall always reckon values from Left's point of view, we call the values of these three positions 1, 2, and 3 in the order given, and the corresponding positions  ,  ,  will have values -1, -2, -3 since in them the free moves are reserved for Right. A position in which neither player has a legal move, for example the position  in Domineering, has the value 0.

Suppose that for some $n = 0, 1, 2, 3, \ldots$, Left has a move to some position of value n, but no other move, and that Right has no legal move whatever. Then of course the value of that position is $n + 1$. In symbols, we write

$$\{n \mid \} = n + 1 \qquad (n = 0, 1, 2, 3, \ldots)$$

indicating (the values of) Left's options before the bar, and Right's after it. It will not affect matters if we give Left some additional moves to positions of integer values less than $n + 1$; so for example

$$\{0, 5, 3, 5 \mid \} = 6.$$

Nor will it matter if Right is allowed to move to some positions whose values are integers, provided these are greater than $n + 1$, so

$$\{1, 4, 7 \mid 13, 20\} = 8.$$

By reversing the roles of Left and Right in this equation, we find also

$$\{-13, -20 \mid -1, -4, -7\} = -8.$$

The simplest formula of this type is of course the equation

$$\{ \mid \} = 0$$

which expresses the fact that positions in which neither player can move have value zero. But here too we can add certain options without affecting the value—namely moves for Left to positions of negative value, or for Right to positions of positive values only. So for example

$$\{-1 \mid 3, 5\} = 0.$$

We can summarize these statements as follows:

When the value of a position is a number, it is the *simplest* number that is neither less than or equal to any of Left's options nor greater than or equal to any of Right's.

We shall call this the *simplicity principle.* The number 0 is the simplest number of all, then come 1 and -1, then 2 and -2, and so on.

Many positions in partizan games have integer values, and a few games consist entirely of such positions, for example:

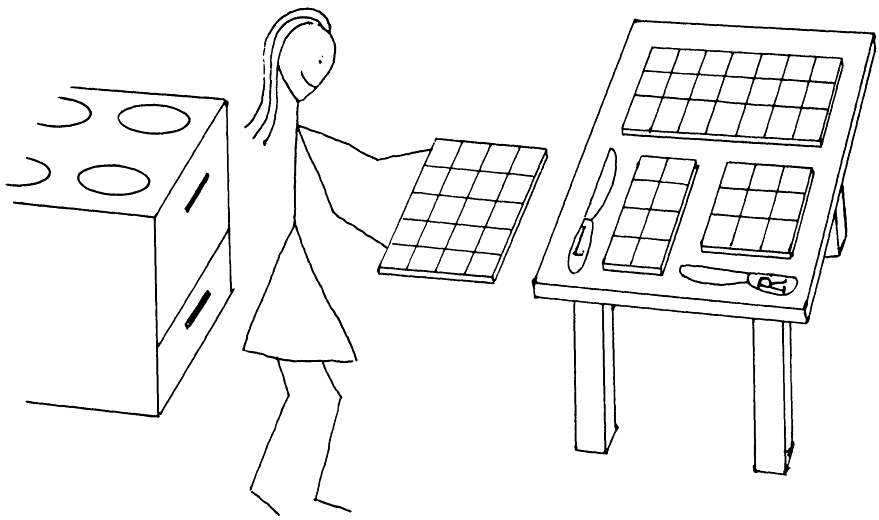


FIG. 4. Ready for a game of Cutcake.

Cutcake is played by two children with a number of rectangular pieces of cake which their mother has already scored along horizontal and vertical lines so as to be ready for breaking into little squares as shown in Figure 4. Lefty moves by breaking some piece along one of the vertical lines, and his sister Rita by breaking a piece along one of the horizontal lines. The values are shown in Table 1. We leave

	1	2	3	4	5	6	7	8	9	10	11	12
1	0	1	2	3	4	5	6	7	8	9	10	11
2	−1	0	0	1	1	2	2	3	3	4	4	5
3	−2	0	0	1	1	2	2	3	3	4	4	5
4	−3	−1	−1	0	0	0	0	1	1	1	1	2
5	−4	−1	−1	0	0	0	0	1	1	1	1	2
6	−5	−2	−2	0	0	0	0	1	1	1	1	2
7	−6	−2	−2	0	0	0	0	1	1	1	1	2
8	−7	−3	−3	−1	−1	−1	−1	0	0	0	0	0
9	−8	−3	−3	−1	−1	−1	−1	0	0	0	0	0
10	−9	−4	−4	−1	−1	−1	−1	0	0	0	0	0

TABLE 1. Values of positions in Cutcake.

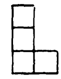
the reader to verify that the pattern indicated by the dividing lines continues indefinitely. The values are computed using only the rules we have given—for example the 5×10 rectangle has the options shown in

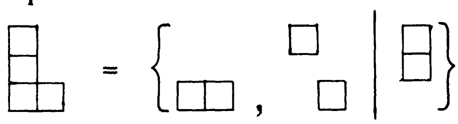
$$\{5\times 1+5\times 9, 5\times 2+5\times 8, 5\times 3+5\times 7,\ldots \mid 1\times 10+4\times 10, 2\times 10+3\times 10\}$$

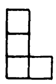
and so has the value $\{-4+1,\ -1+1,\ -1+0, 0+0, 0+0 \mid 4+1, 4+4\}$ or simply

$$\{-3, 0, -1, 0, 0 \mid 5, 8\} = 1,$$

by the simplicity rule.

Fractional Values. When we try to evaluate the position  in Domineering, we find, in the symbolic notation, the equation

 $=\{-1, 0 \mid 1\}$

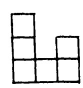
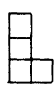
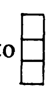
and unfortunately there is no *integer* strictly between -1 and 0 (on the left), and 1 (on the right). However, there is a *fraction*, and in fact the simplest such fraction is $\frac{1}{2}$. It turns out that in fact the position  is really worth exactly $\frac{1}{2}$ a move to Left, in a suitable sense, and that adding two such regions to a Domineering position confers exactly the same advantage on Left as adding  would.

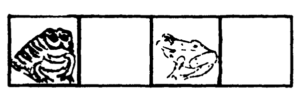
Other positions in finite games can have values involving quarters or eighths of moves etc., and the simplicity rule continues to hold for such values, provided we add the conditions that all integers are simpler than fractions with denominator 2, while these are simpler than those with denominator 4, in turn simpler than those with denominator 8, and so on. So for example

$$\{1 \mid 1\tfrac{3}{8}\} = 1\tfrac{1}{4}$$

since there is no integer or half-integer between 1 and $1\frac{3}{8}$, but $1\frac{1}{4}$ *is* between these two numbers. Other examples are

$$\{-1 \mid 21\tfrac{1}{2}\} = 0,\ \{0 \mid 3\tfrac{1}{4}\} = 1,\ \{\tfrac{1}{4} \mid 2\} = 1,\ \{\tfrac{1}{4} \mid 1\} = \tfrac{1}{2}.$$

Fractional positions arise in many games. In Domineering the position  has value $\frac{3}{4}$, since Left's best option is to  (value $\frac{1}{2}$), and Right's is to   (value 1), and we have $\{\frac{1}{2} \mid 1\} = \frac{3}{4}$.

In Toads and Frogs the position 

turns out to have value $\frac{1}{2}$. The first of these examples is rather tricky to evaluate, since one of Left's options does not have a numerical value, but the second example is easy, and the reader will probably benefit by verifying that all the values shown in Figure 5 are consequences of the simplicity rule.

In **Red-blue Hackenbush** every position has a numerical value, which may be integral or fractional. This game is played with *pictures* made of red and blue edges, and every edge must touch the ground or be connected to the ground by some chain of other edges. Left moves by chopping some bLue edge, and Right by chopping a Red one, and after each chop, any edges no longer connected to the ground are deleted.

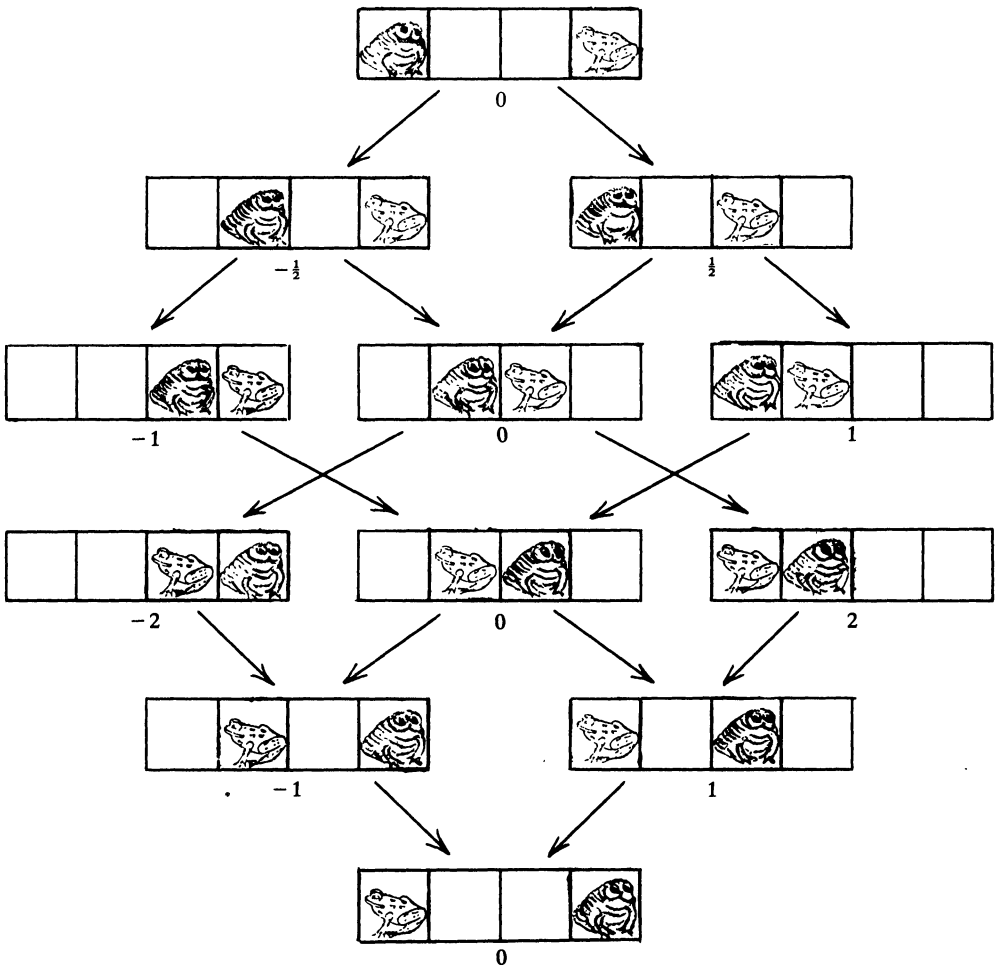
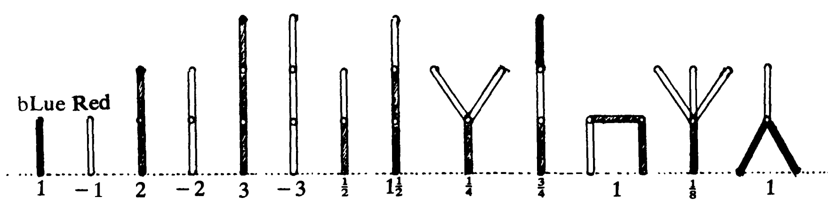
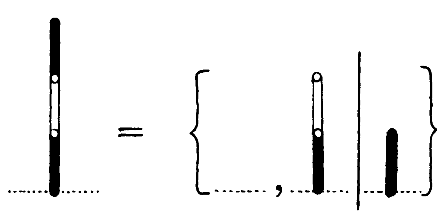


FIG. 5. Anatomy of 4-place Toads and Frogs.

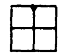
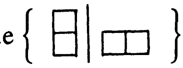
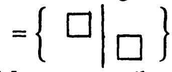
Exercises:



Example:

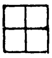
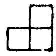


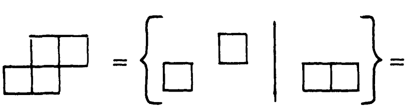
$= \{0, \frac{1}{2} \mid 1\} = \frac{3}{4}.$

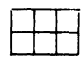
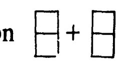
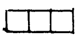
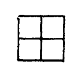
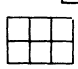
Fuzzy values. In Domineering, the position  has value  $= \{1 | -1\}$, which is not a number, since Left's option is greater than Right's. We call this value ± 1, and in general use $\pm x$ for $\{x | -x\}$, which for $x \geqq 0$ is a non-numerical value. The position  $=$  has value ± 0, which we also call $*$, and which is a very common value indeed. More generally, there is a non-numerical value $\{x \mid y\}$ for every pair of numbers x and y with $x \geqq y$. In a sense soon to be explained, the value $\{x \mid y\}$ is *strictly less* than every number greater than x, *strictly greater* than every number less than y, but *incomparable* with all numbers between x and y inclusive. There is a simple policy for finding the best move from any sum of such values, possibly together with numerical values:

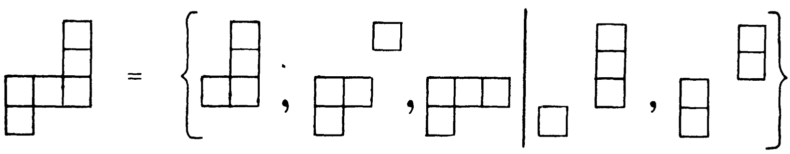
Never move in a component whose value is a number unless you have no other alternative. Of the various components $\{x \mid y\}$ with $x \geqq y$, move in one with the largest possible value of $x - y$.

Since $\frac{1}{2}(x - y)$ is called the *temperature* of $\{x \mid y\}$, the policy may be summarized more briefly: move in the *hottest* $\{x \mid y\}$. If several components are equally hot, it will not matter which of them we choose. A similar temperature policy applies in many other situations, but not in all.

We can now analyze the position of Figure 1. We have already met the values $\{1 \mid -1\}$ and $\{0 \mid 0\}$ of  and , and the evaluation

 $= \{0 \mid -1\}$

is easy. The region  is a little harder. Left has the option  $= 2$, and all Right's options are similar to  $= -\frac{1}{2}$. But Left has also the option , of value ± 1. But since $\pm 1 < 2$, this is worse than his previous option, and so we have  $= \{2 \mid -\frac{1}{2}\}$. Finally, we have the equation



showing that this region has value $\{\frac{1}{2}, *, -\frac{1}{2} | 1, 2\}$. Here since $\frac{1}{2}$ is greater than $*$ and $-\frac{1}{2}$, and 1 is less than 2, the value is $\{\frac{1}{2} | 1\} = \frac{3}{4}$.

Since we can neglect the isolated squares, which neither player can use, this calculation yields

$$\{2 \mid -\tfrac{1}{2}\} + \{1 \mid -1\} + \{0 \mid -1\} + \{0 \mid 0\} + \tfrac{3}{4}$$

for the total value of Figure 1, where we have kindly rearranged the terms $\{x \mid y\}$ in decreasing order of their temperatures $\frac{1}{2}(x - y)$. If both players play according to the temperature policy, the value after the first four moves will be

$$2 - 1 + 0 + 0 + \tfrac{3}{4} = 1\tfrac{3}{4}$$

if Left starts, and

$$-\tfrac{1}{2} + 1 - 1 + 0 + \tfrac{3}{4} = \tfrac{1}{4}$$

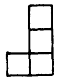
if Right starts. Since both values are positive, Left can win no matter who starts.

If however, Left moves first and makes his move rather stupidly in the top right hand corner, this has the effect of replacing $\{2 \mid -\frac{1}{2}\}$ by ± 1, and making the total value

$$\{1 \mid -1\} + \{1 \mid -1\} + \{0 \mid -1\} + \{0 \mid 0\} + \tfrac{3}{4}$$

and if the next four moves are played sensibly, the resulting value will be

$$-1+1-1+0+\tfrac{3}{4}=-\tfrac{1}{4},$$

and since this is negative, Right will win. If instead Left's opening move had been in the bottom left-hand corner, replacing $\frac{3}{4}$ by  $=\frac{1}{2}$, the value would have become

$$\{2\mid-\tfrac{1}{2}\}+\{1\mid-1\}+\{0\mid-1\}+\{0\mid 0\}+\tfrac{1}{2},$$

and after four sensible moves, which lead to the value

$$-\tfrac{1}{2}+1-1+0+\tfrac{1}{2}=0,$$

we can pretend that the game has finished. Since the most recent move here was Left's, he has won, but it was a close shave, and another false step would have been disastrous.

In general, the winner, assuming best play, can be determined from the value according to the following rule:

> If the value is *positive*, *Left* can win, no matter who starts.
> If the value is *negative*, *Right* can win, no matter who starts.
> If the value is *zero*, whoever plays *second* can win.
> If the value is *fuzzy*, the *first* player to move can win.

The *fuzzy* games are those that are neither *positive*, *negative*, nor *zero*, but rather, *confused* with zero. Such a game is $\{x\mid y\}$, when $x\geqq 0\geqq y$.

Comparing and adding values in general. We write G^L for the typical option available to Left from G, and G^R for the typical option of Right, so that symbolically

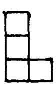
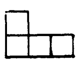
$$G=\{G^L\mid G^R\}.$$

This notation should not be taken to imply either the uniqueness or even the existence, of the options for either player. So for instance from the game $G=\{1,\pm 2\mid\ \}$, from which Left can move to options of value either 1 or ± 2, but Right cannot move, G^L means either 1 or ± 2, but G^R takes no value (which is different from saying that it takes the value 0). In this notation, the sum of two games is given by the formula:

$$G+H=\{G^L+H, G+H^L\mid G^R+H, G+H^R\}.$$

We can also define the *negative* of a game by the formula

$$-G=\{-G^R\mid -G^L\},$$

which merely expresses the fact that the roles of Left and Right are to be reversed throughout. Thus the negative of  is , and the negative of the Hackenbush position  is 

The nomenclature is justified by the following theorems:

1. *Adding a zero game (that is, one won by the second player), never affects the outcome under best play.*
2. *The sum of any game and its negative is zero.*
3. *The sum of two games that are both positive or zero is itself positive or zero.*
4. *The sum of two negative or zero games is another of the same kind.*
5. *If the difference $G-H$, by which we mean $G+(-H)$, is zero, we can replace G by H in any sum without affecting the outcome of that sum.*

6. *If $G \geqq H$, by which we mean that $G - H \geqq 0$, then Left need not object when we replace H by G in any sum, and Right need not object to the replacement of G by H.*

We can use these results to justify the formal definition of value. If $G - H$ is a zero game, then we say that G and H have the same *value*, and write $G = H$. In general order relations between games and their values can be decided by the condition that ? has the same value in the relation G ? H as it does in $G - H$? 0. We always have one of the four cases

$G > H$, meaning that $G - H$ can always be won by *Left*
$G < H$, meaning that $G - H$ can always be won by *Right*
$G = H$, meaning that $G - H$ can be won by the *second* player
$G \parallel H$, meaning that $G - H$ can be won by the *first* player.

We also abbreviate various compounds of these relations in natural ways. Thus $G \geqq H$ means $G > H$ *or* $G = H$, $G \lhd\!\!| H$ means $G < H$ *or* $G \parallel H$.

We can use these conditions to check our assertions about values $\{x \mid y\}$ when x and y are numbers with $x \geqq y$. For if z is any number in the range $x \geqq z \geqq y$, then from the difference $\{x \mid y\} - z$, Left can win by moving to $x - z \geqq 0$, and Right by moving to $y - z \leqq 0$. On the other hand, if $z > x$, then $\{x \mid z\}$ is some number between x and z, so we have $\{x \mid y\} \leqq \{x \mid z\} < z$.

In checking inequalities between values we can use some other obvious remarks—the value of a game G will be unaltered or increased if we increase the value of any Left or Right option of G, add a new Left option, or remove some Right option.

Dominated and reversible options. In this section, we sketch a method by which one can reduce any value to a simplest form. Two games have the same value if and only if their simplest forms are identical. We describe some modifications to the form of a game which do not affect its value.

If $A \leqq B$, and A and B are both Left options from G, then since Left will in any case prefer his move to B over his move to A, the value of G will not be affected if we omit A, while retaining B. If A and B were Right options we could instead have omitted B and retained A. In each case, we say that the option we intend to omit is *dominated* by the retained one.

A more subtle concept is that of *reversible* options. Suppose that the Left option G^{L_0} of G has itself a Right option $G^{L_0R_0} \leqq G$. Then in turns out that the value of G is unaffected if we delete the

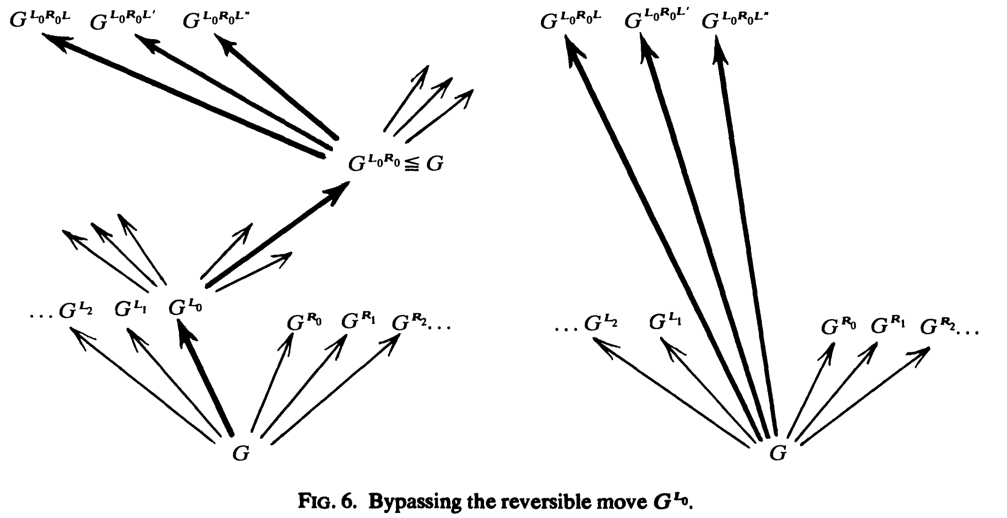


FIG. 6. Bypassing the reversible move G^{L_0}.

G^{L_0} as a Left option of G, and insert instead all the Left options $G^{L_0R_0L}$ of $G^{L_0R_0}$ as new Left options of G. In this case we say that the option G^{L_0} was reversible *through* $G^{L_0R_0}$ *to* the $G^{L_0R_0L}$, and we call this process the *bypassing* of G^{L_0}. See Figure 6. Of course we call the Right option G^{R_0} reversible if it has some Left option $G^{R_0L_0} \geqq G$, and we can then replace it by all the $G^{R_0L_0R}$ as new Right options of G. In general, an option is reversible if the opponent can move from it to a position which is better for him than the original game G.

Here is a simple example. The game $G = \{0, \pm 1 \mid 2\}$ plainly has no dominated options, since 0 and ± 1 are incomparable. But if Left moves to ± 1, Right's reply to -1 is obviously better for Right than the original game, which was clearly positive. So the option ± 1 is reversible through -1, and we can replace it as a Left option of G by the list of all Left options of -1 without affecting the value of G. Since in fact -1 has no Left options, we find $G = \{0 \mid 2\} = 1$.

If we eliminate all dominated and reversible options from all positions of some game G with finitely many positions, we finally obtain *the simplest form of* G. The following result justifies the name:

> Two games G and H have the same value if and only if their simplest forms are identical.

There is also a slightly weaker result valid for infinite games, which we do not discuss here.

Impartial games and the game of Nim. A game is called *impartial* if from every position the moves available to the two players are exactly the same. The theory for such games was described by R. P. Sprague [**10**], and independently by P. M. Grundy [**6**] and seems to have been independently rediscovered several times since then. It fits very naturally inside our more general theory.

The game of Nim (Bouton [**2**]) is played with a number of heaps, of beans say, and the legal move, for either player, is to reduce the size of some heap by removing some of its beans. We write $*n$ for the value of a Nim-heap of size n, so that

$$*n = \{*0, *1, \ldots, *(n-1) \mid *0, *1, \ldots, *(n-1)\}.$$

The three simplest cases are

$$*0 = 0, *1 = *, \quad \text{and} \quad *2 = \{0, * \mid 0, *\}.$$

Then the Sprague–Grundy theory is contained in the following assertions:

(i) Every impartial game with only finitely many positions has one of the values $*0, *1, *2, \ldots$.

(ii) $\{*a, *b, *c, \ldots \mid *a, *b, *c, \ldots\} = *m$, where m is the least number from $0, 1, 2, \ldots$ that does *not* appear among $a, b, c, \ldots$ (m is called the *mex* of $a, b, c, \ldots$).

(iii) If $a, b, c, \ldots$ are distinct, we have

$$*2^a + *2^b + *2^c + \cdots = *(2^a + 2^b + 2^c + \cdots).$$

(iv) $*n + *n = 0 (= *0)$.

So for example, from property (ii) we have $\{*0, *1, *4, *7 \mid *0, *1, *4, *7\} = *2$ and, as an example of addition:

$$*3 + *5 = (*2 + *1) + (*4 + *1) = *2 + *4 = *6$$

using properties (iii) and (iv).

These results are in fact very easy to prove, for we can see that any $*n$ with $n > m$ will be a reversible option in

$$*m = \{*0, *1, \ldots, *(m-1) \mid *0, *1, \ldots, *(m-1)\}$$

which proves (ii), from which (i) follows inductively, and establishes that $*x + *y$ is equal in value to *some* Nim-heap $*z$, and then the rules of (iii) and (iv) for evaluating z are easily established. The

whole theory generalizes trivially to infinite impartial games, there being a value $*\alpha$ for every ordinal number α.

The game of Kayles. Grundy's game. Figure 7 shows a position in the old English bowling game of Kayles, introduced by Dudeney [5]. The players alternately remove either *any one*, or *any two adjacent* skittles, and the first who is unable to do so is the loser.

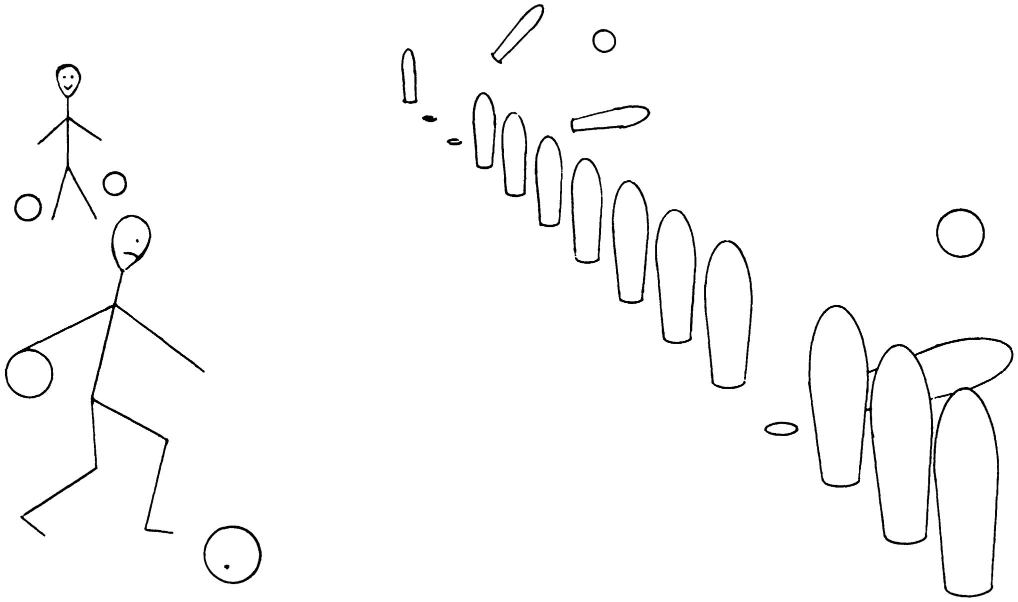


FIG. 7. A game of Kayles.

Kayles is obviously an impartial game, and if we write K_n for the value of a line of n skittles in Kayles, we have

$$K_n = \{K_a + K_b \mid K_a + K_b\} \quad a + b = n - 1 \text{ or } n - 2.$$

Thus for example, neglecting $K_0 = 0$, we find

$$K_5 = \{K_4, K_3 + K_1, K_2 + K_2, K_3, K_2 + K_1 \mid \text{ditto}\},$$

and using the values $*0, *1, *2, *3, *1$ for K_0, K_1, K_2, K_3, K_4 that we can suppose ourselves to have already found, this becomes

$$\{*1, *3 + *1, *2 + *2, *3, *2 + *1 \mid \text{ditto}\}$$

$$= \{*1, *2, *0, *3, *3 \text{ ditto}\} = *4, \text{by the mex rule.}$$

R. K. Guy discovered in this way the remarkable fact that the sequence of values K_n in Kayles is periodic with period 12 for $n \geqq 71$, and many similar results have been found for other games with heaps (Guy and Smith [7]). *Grundy's game* (split any heap into two smaller heaps of distinct sizes) has been analyzed by Berlekamp to 240,000 values and has not yet become periodic. However, a number of structural features were observed in the values, and if these persist we can prove that the values *will* ultimately become periodic though perhaps only for much larger n.

Seating couples and larger families. Superstars. Figure 8 shows the dinner party which celebrates the end of one of the chapters of our forthcoming book. Left and Right are responsible for the seating

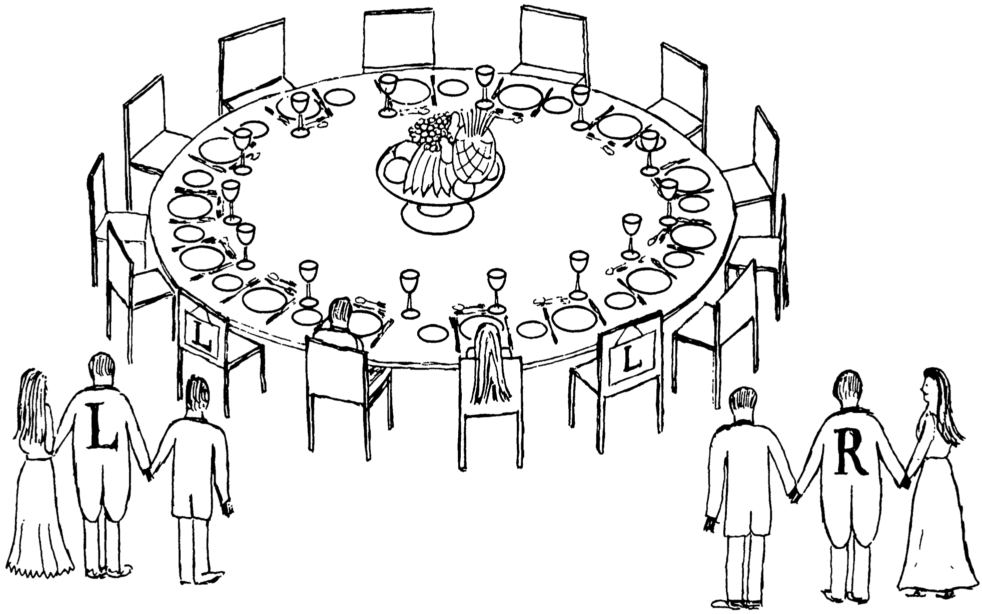


FIG. 8. Seating Couples.

arrangements, and will alternately seat each pair of guests as they arrive. Left thinks it proper to seat a lady only to the *left* of her partner, while Right prefers to seat her to the *right.* No lady may be seated next to a gentleman who is not her partner. Whoever is first unable to seat a couple according to these self-imposed rules has the embarrassing task of explaining the situation to the remaining guests, and may be said to *lose.*

In this game, whenever one of the two players seats a couple, he effectively reserves the two adjacent seats for use of his opponent, since neither player can seat two couples in four consecutive seats. So after the first move, each line of n seats is terminated by reserved seats at either end, and so has the form

LnL, if the two end seats are reserved for *Left*
RnR, if they are reserved for *Right,* and
LnR or RnL if they are reserved one for each player.

We have the equations:

$$LnL\,(=-RnR)=\{LaR+RbL \mid LbL+LaL\}$$

$$LnR\,(=RnL)=\{LaR+RbR \mid LbL+LaR\},$$

where a and b range over all non-negative integers with $a+b=n-2$, except that the number x in any position LxR or RxL must be strictly positive.

There is a generalization of the impartial theory which is very useful in analyzing such games, namely:

> If a game of the form
>
> $$\{A, B, C, \ldots \mid -A, -B, -C, \ldots\}$$
>
> has the properties $A \leqq -A$, $B \leqq -B$, $C \leqq -C, \ldots$, then it has value $*m$, where m is the least number, from $0,1,2,3,\ldots$, for which $*m$ does not appear among the values $A, B, C, \ldots$.

Since we can see easily that $RbR \leqq LbL$, this theorem shows in particular that each LnR has a value of form $*m$, where m is the least number for which $*m$ is not an option of LnR. The values that arise in Seating Couples are very simple, and the reader might care to work them out himself as an exercise.

We may modify the game by considering families of size $n > 2$ instead of couples. For example in the game of Seating Families of Five, each family consists of a Mother, a Father, and three children who must always be seated between their parents. Once again, Left will seat each lady at the left end of her family, while Right prefers to seat her to the right, and no lady may be seated next to the husband of another. The values are then given by the same equations as for our game of seating couples, except that we have $a + b = n - 5$, rather than $n - 2$.

We mention this particular example because of the rather special behavior of its values. It can be shown that the sequence

$$L0R, L1R, L2R, \ldots$$

for Seating Families of Five, consists of the values for the (impartial) game of *Dawson's Kayles* (defined like ordinary Kayles, except that the *only* move is to remove two adjacent skittles), each repeated three times. Moreover, the values LnL and RnR usually have the form

$$G = \{*a, *b, *c, \ldots \mid *A, *B, *C, \ldots\}.$$

How do we analyze these?

It turns out that the value of the above game depends rather critically on the two numbers

$$m = \operatorname{mex}(a, b, c, \ldots) \text{ and } M = \operatorname{mex}(A, B, C, \ldots).$$

If $m = M = n$, say, we have $G = *n$. If, however, $m > M$, it turns out that the value of G is otherwise independent of the particular numbers $a, b, c, \ldots$, and we write it as

$$\uparrow_{ABC\ldots} = \{*0, *1, *2, \ldots \mid *A, *B, *C, \ldots\}$$

where the lefthand side of the bracket may be thought of as containing $*n$ for *all* integers n. If $m < M$, then of course G has value

$$\downarrow^{abc\ldots} = \{*a, *b, *c, \ldots \mid *0, *1, *2, \ldots\}$$

which is the negative of $\uparrow_{abc\ldots}$.

In a sense to be discussed later, we say that $\uparrow_{ABC\ldots}$ has *atomic weight* $+1$ and $\downarrow^{ABC\ldots}$ atomic weight -1, while $*n$ has atomic weight 0. For sums of these games it is not hard to see that if the total atomic weight is 2 or more Left can always win, and he can also win if the atomic weight is 1 and he has the starting move. Moreover, we have the important *translation property*:

> If the values $*A, *B, *C, \ldots$ are merely $*a + *n, *b + *n, *c + *n, \ldots$ in some order, for *some n*, then we have the equation $\uparrow_{ABC\ldots} = \uparrow_{abc\ldots} + *n$, *for the least such n.*

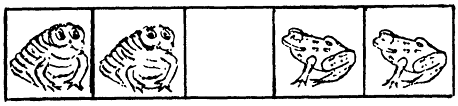
The particular case $\uparrow_1 = \{0 \mid *\}$ in simplest form is usually called $\uparrow$, and pronounced 'up', with negative $\downarrow$, pronounced 'down'. From the translation property, and the formulas for $*a + *b$, we have also the equations

$\uparrow_0 = \uparrow + * = \{0, * \mid 0\}$ in simplest form
$\uparrow_2 = \uparrow + *3 = \{0 \mid *2\}$ in simplest form
$\uparrow_3 = \uparrow + *2 = \{0 \mid *3\}$ in simplest form, and so on.

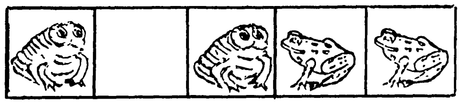
The value $\uparrow$ is interesting because it is strictly positive, and so in favor of Left, but it is also strictly less than every positive number, even than every positive infinitesimal number. In the sense of [3] we may call it *small.*

Linear combinations of $\uparrow$ and $*$ arise in many games. For example in Toads and Frogs the starting

position



in a lane of length five, has value $*$, while the position



which arises after one move from this, has value $\uparrow$. A rather unexpected equality is the equation

$$\{0 \mid \uparrow\} = \uparrow + \uparrow + *,$$

which has been called *the upstart identity.*

Thermography and the Mean Value Theorem. We already know that the general game G need not have a numerical value. However, it turns out that there is a best numerical approximation m to G, which has the property that when we play the sum of many copies of G, the result is approximately equal to the same number of copies of the number m, which is called the *mean value* of G. For instance, the game $G = \{\{7 \mid 5\} \mid \{4 \mid 1\}\}$ has mean value $4\frac{1}{4}$, and temperature $1\frac{3}{4}$, and this statement allows us to say that $1000G$ lies between any number strictly less than $4250 - 1\frac{3}{4}$ and any number strictly greater than $4250 + 1\frac{3}{4}$.

The *thermograph* is a device for calculating mean values and temperatures. We draw the number scale as the horizontal axis, but with positive numbers to the left and negative ones to the right, instead of the more usual opposite convention. We then draw the thermographs, supposed already computed, of all the options G^L and G^R of G. The thermograph of a number x is a vertical straight line originating at the point x of the axis.

We then, at the height corresponding to any *temperature*, t, take the Leftmost Right boundary of any G^L, and move it a distance t (equal to the height above the axis) to the *right*, and the Rightmost Left boundary of any G^R, and move it distance t to the *left*. The resulting curves define the boundaries of the thermograph of G itself until they meet, at a height t_0 (called the temperature of G) above the axis, the thermograph above t_0 being a single vertical line called the *mast*.

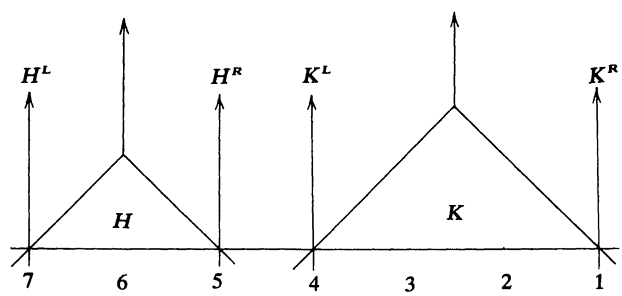


FIG. 9. Thermographs of $H = \{7 \mid 5\}$ and $K = \{4 \mid 1\}$.

A few examples will make the process much clearer (see Figure 9). For the game $H = \{7 \mid 5\}$, the thermograph of $H^L = 7$ is a vertical line through the point 7 on the axis, and that of H^R is a similar vertical through the point 5. So the Left boundary of the thermograph of H starts as a line slanting diagonally up and right through the point 7, and the Right boundary starts as a similar line through 5, slanting diagonally up and left. These meet at a point whose height is 1 above the point 6, and so the temperature of H is 1, and it has mean value 6. Similarly the game $K = \{4 \mid 1\}$ has mean value $2\frac{1}{2}$, and temperature $1\frac{1}{2}$.

Now for the game $G=\{H \mid K\}$, the Right boundary of H starts at 5 and slants up and left to the point 1 above 6, and is vertical thereafter. The Left boundary starts at 4 and slopes diagonally up and right to a point at height $1\frac{1}{2}$ above the point $4\frac{1}{4}$, and is vertical thereafter. The Left and Right boundaries of G, obtained by pushing these inward towards each other, will therefore be a line starting vertically at 5 before turning to slant up and right at height 1, and another starting vertically at 4 before turning to slant up and left at height $1\frac{1}{2}$, as in Figure 10. These meet at height $1\frac{3}{4}$ above the

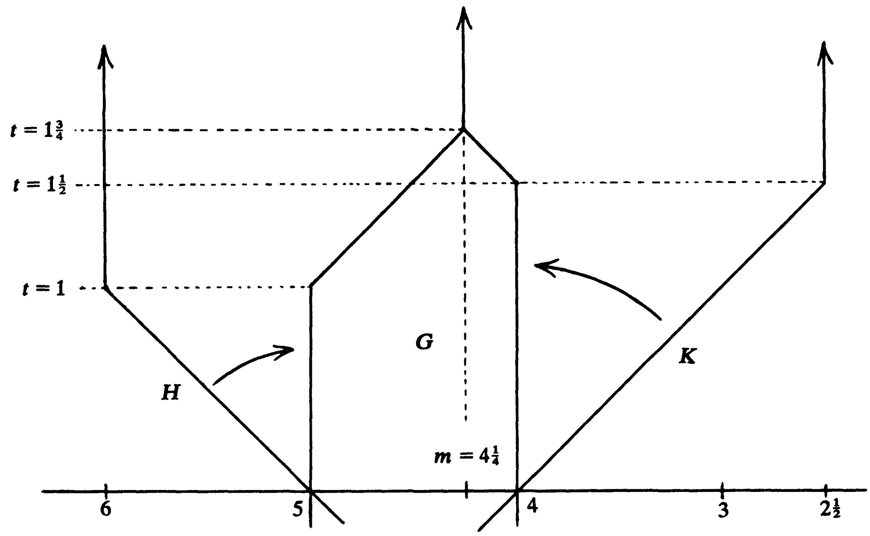


FIG. 10. How to find thermographs.

point $4\frac{1}{4}$ on the axis, showing that the temperature of G is $1\frac{3}{4}$ and its mean value $4\frac{1}{4}$. The same algorithm works for any G for which the Leftmost right boundary of any G^L starts to the left of the Rightmost left boundary of any G^R—in the other cases G is a number, namely that given by the simplicity principle. If t' is any number greater than the temperature of G and m is the mean value of G, we have

$$nm-t'<n\cdot G<nm+t'$$

for any positive integer n, justifying the terminology.

More generally, if games $G_0, G_1, G_2, \ldots$ have the respective mean values $m_0, m_1, m_2, \ldots$, then we can say that

$$m_0+m_1+m_2+\cdots-t'<G_0+G_1+G_2+\cdots<m_0+m_1+m_2+\cdots+t'$$

for any number t' greater than the temperature of every G_i. This often helps us to win by showing that a complicated sum of games is positive without making a detailed analysis.

Atomic weights. For small games like $\uparrow$ we need a more delicate scale (both mean value and temperature of $\uparrow$ are 0) This is provided by the **atomic weight calculus**, which obtains for such small games some of the benefits that thermography has given for large ones. The details are unexpectedly subtle.

There is an atomic weight G'' defined for any game G for which no position has value a non-zero number. This may be computed by the formula

$$G''=\{(G^L)''-2 \mid (G^R)''+2\},$$

except that when this defines an integer, it might not be the correct integer. The correct integer in such a case is

the *largest* n with $(G^L)'' - 2 \lhd| n \lhd| (G^R)'' + 2$, if G exceeds remote stars,
the *least* such n if G is exceeded by remote stars, and
the integer *zero*, if G is incomparable with remote stars.

The *remote stars* for a game G are those Nim-heaps $*N$ which do *not* occur as values of positions of G. It can be shown that G has the same order-relations with all the remote stars for G.

As an example of the atomic weight calculus, we take $G = \{0 \mid \uparrow\}$. Since 0 has atomic weight 0 and $\uparrow$ atomic weight 1, the formula would give

$$G'' = \{0 - 2 \mid 1 + 2\} = \{-2 \mid 3\}$$

which defines the integer 0. Since this is an integer, we cannot immediately assert that it is the atomic weight, but must first compare G with the remote stars. The smallest star that is remote for G is $*2 = \{0, * \mid 0, *\}$, and we find, by playing the game, that $G + *2$ is positive, so that G exceeds the remote stars and so is the greatest integer $\lhd| 3$, namely 2. Of course the value 2 found this way for the atomic weight is consistent with our earlier evaluation of this G as $\uparrow + \uparrow + *$.

As another example, we take $H = \{\uparrow + \uparrow + * \mid \downarrow + *\}$. The atomic weight formula now gives $\{2 - 2 \mid -1 + 2\} = \{0 \mid 1\} = \frac{1}{2}$, which therefore really is the atomic weight. In fact this game is, in a definite sense, one-half of $\uparrow$, and is usually written $\frac{1}{2} \cdot \uparrow$. We can define $x \cdot \uparrow$ for all *non-integral* x by the formula

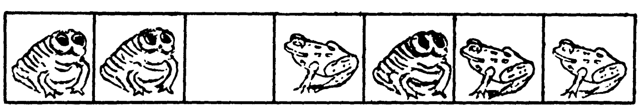
$$x \cdot \uparrow = \{(x^L + 2) \cdot \uparrow + * \mid (x^R - 2) \cdot \uparrow + *\},$$

and this, taken together with the obvious definition for integral x, satisfies the distributive law

$$(x + y) \cdot \uparrow = (x \cdot \uparrow) + (y \cdot \uparrow).$$

The particular cases $\frac{1}{2} \cdot \uparrow$, $\frac{1}{4} \cdot \uparrow$, $* \cdot \uparrow$ arise in the twisted form of Bynum's game [**4**, 199–200] which we do not have time to describe here. An even more interesting sequence of infinitesimals turns up in the untwisted form. The theory of these and many other games is made much easier by the most useful property of atomic weights asserting that if a game has atomic weight 2 or more, it is positive.

For example the Toads and Frogs position



can be computed to have the value $\{\Uparrow * \mid 0\}$ (where $\Uparrow * = \uparrow + \uparrow + *$) whose atomic weight is $\{2 - 2 \mid 0 + 2\} = 1$. So the position of Figure 11 which has atomic weight 2, must be positive and Left

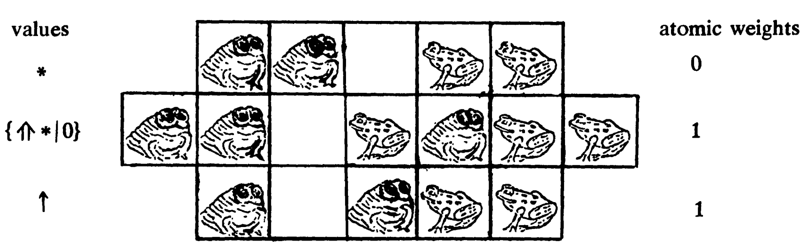


FIG. 11. A Toads and Frogs position of Atomic Weight 2.

should win no matter who starts. In general Left can be sure of winning any game whose atomic weight is 2 or more, but if he has the move, he can win provided only that the atomic weight is fuzzy or positive. (Atomic weights *can* be fuzzy, for example $\{\Uparrow \mid \Downarrow\}$ has atomic weight $\{2 - 2 \mid -2 + 2\} = *$, but fortunately they are usually integers.)

The work of Milnor and Hanner. This theory is related to that expounded by John Milnor [**9**]. Although there are a number of similarities, there are also some important differences. Milnor introduces numbers in an *ad hoc* fashion, as *payoffs*, but for us they arise naturally from the theory. Finite games lead only to the dyadic rationals, but since we allow infinite games satisfying the finishing condition we do in fact get all real numbers as well as the vast array of infinite and infinitesimal numbers like

$$\omega, \omega+1, \omega-1, \omega/2, \sqrt{\omega}, 1/\omega, 1/\omega^2, 1/\omega^\omega,$$

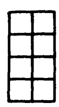
of [**3**].

Milnor's *incentive* corresponds roughly to our *temperature* and Olof Hanner [**8**] has given a rather complicated proof of the Mean Value Theorem for Milnor's games. But for Milnor the *only* games of incentive zero are the real numbers, so he has no analogs of our games

$$*, *2, \uparrow, \text{etc.}$$

The games covered by his theory are just those in which every position either has positive temperature or is a *real* number.

We close with a few valedictory remarks. Firstly it seems that the logically most natural theory is that allowing infinite games, subject to the finishing condition. Most of the theory makes no distinction between the finite and infinite case, two notable exceptions being the mean value theorem and Simon Norton's theorem that games with only finitely many positions cannot have odd additive order. (Norton has constructed infinite games of all orders.) Seen in this light, the theory includes the theory of infinite and infinitesimal numbers developed in [**3**], and greatly extends that theory.

The world of all games has many interesting highways and byways of which the examples here (numbers, values $\{x \mid y\}$, values $*n$, values $\uparrow_{abc\ldots}$) are only a few. There are powers of $\uparrow$ (for example $\uparrow^2 = \{0 \mid \downarrow *\}$) as well as multiples of $\uparrow$, and some extremely small values, for example the value $+_2$ (pronounced "tiny-two") $= \{0 \mid \{0 \mid -2\}\}$ of the Domineering position . Fortunately, it seems that the most illuminating way to examine this vast wealth of structure is to investigate various partizan games that are naturally suggested as interesting to play. Why don't *you* help to investigate these games and their values, weird and wonderful? It's good fun!

This paper was written in haste. That I have not repented it at leisure is entirely due to the efforts of Richard Guy and Karen McDermid.

References

1. E. R. Berlekamp, J. H. Conway and Richard K. Guy, Winning Ways, Freeman, San Francisco, 1977.
2. C. L. Bouton, Nim, a game with a complete mathematical theory, Ann. of Math., Princeton (2) 3(1902) 35–39.
3. J. H. Conway, All numbers great and small, Univ. of Calgary Math. Dept. Research Paper No. 149, February, 1972.
4. ———, On Numbers and Games, Academic Press, New York, 1976.
5. H. E. Dudeney, Canterbury Puzzles, London, 1910, 118, 220.
6. P. M. Grundy, Mathematics and games, Eureka, 2 (1939) 6–8.
7. Richard K. Guy and Cedric A. B. Smith, The G-values of various games, Proc. Cambridge Philos. Soc., 52 (1956) 514–526.
8. Olof Hanner, Mean play of sums of positional games, Pacific J. Math., 9 (1959) 81–99; MR 21 # 3277.
9. John Milnor, Sums of positional games, in H. W. Kuhn and A. W. Tucker (eds.) Contributions to Theory of Games II, Ann. of Math. Studies # 28, Princeton, 1953, 291–301; MR 14–779.
10. R. P. Sprague, Über mathematische Kampfspiele, Tôhoku Math. J., 41 (1935–36) 438–444.

DEPARTMENT OF MATHEMATICS AND MATHEMATICAL STATISTICS, UNIVERSITY OF CAMBRIDGE, 16 MILL LANE, CAMBRIDGE CB2 1SB, ENGLAND.

SOME EXTENSIONS OF THE TIETZE–URYSOHN THEOREM

E. K. VAN DOUWEN, D. J. LUTZER, AND T. C. PRZYMUSIŃSKI[1]

Abstract: In this note we examine certain sharpenings of the Tietze–Urysohn extension theorem. We prove, for example, that if A is a closed subspace of a normal space X then there is a ***continuous*** extender $\eta: C^*(A) \to C^*(X)$, where $C^*(A)$ and $C^*(X)$ are the Banach spaces of continuous, bounded, real-valued functions on A and X, respectively. However, well-known examples show that we cannot arrange that η is both continuous and linear, or that η is an isometry.

Introduction. The classical Tietze–Urysohn theorem guarantees that every continuous, bounded, real-valued function f defined on a closed subspace A of a normal space X can be extended to a continuous, real-valued function $\hat{f}$ defined on all of X and having the same bounds as f. If we let $C^*(A)$ and $C^*(X)$ denote the real vector spaces of continuous, bounded, real-valued functions on A and X respectively, then the Axiom of Choice yields the following slight reformulation of the classical theorem.

A. THEOREM (Tietze–Urysohn). *If A is a closed subspace of the normal space X then there is a function $\eta: C^*(A) \to C^*(X)$ such that for every $f \in C^*(A)$, $\eta(f)$ extends f and has the same bounds as f.*

Any function $\eta: C^*(A) \to C^*(X)$ having the property that $\eta(f)$ extends f whenever $f \in C^*(A)$ will be called an *extender* from $C^*(A)$ to $C^*(X)$. It is natural to wonder whether, if one selects the extension $\eta(f)$ of f more carefully, one can obtain an extender η which respects the natural linear and/or topological structures of the function spaces in Theorem A. For metric spaces, an affirmative answer is provided by the Dugundji Extension Theorem.

B. THEOREM [D]. *Suppose A is a closed subset of a metrizable space X. Then there is a function $\eta: C^*(A) \to C^*(X)$ such that:*

(1) *for each $f \in C^*(A)$, $\eta(f)$ extends f and has the same bounds as f;*

(2) *the function η is linear.*[2]

Theorem B is a model for all of the extension theorems considered in this paper so we shall pause to make a few comments about its conclusions and its proof.

In introducing the theorem we said that it shows that η can be constructed in such a way that η respects the natural topological structures of the function spaces. Let us now introduce one such structure, the sup-norm topology. For a function $f \in C^*(A)$ the *sup-norm of f* is defined by $\|f\| = \sup\{|f(x)| : x \in A\}$; the norm of a function $F \in C^*(X)$ is analogously defined. When equipped with these norms, $C^*(A)$ and $C^*(X)$ are Banach spaces and the two conclusions of Theorem B force η to be an isometry. That may be seen as follows. Let $f, g \in C^*(A)$. Then $\|\eta(f) - \eta(g)\| = \|\eta(f-g)\|$ since η is linear, and because $\eta(f-g)$ has the same bounds as $f-g$, $\|\eta(f-g)\| = \|f-g\|$. Thus η is a linear, isometric extender.

The proof of Theorem B is too technical to include here but its central idea is easily described, at least in principle. One does not extend the functions in $C^*(A)$ one at a time as in the proof of Theorem A; instead one defines an extension process which acts simultaneously (and linearly) on all members of $C^*(A)$. Dugundji's extension process has been improved in recent years so that it now

1. This paper was completed while the third author was visiting the University of Pittsburgh as a Mellon Postdoctoral Fellow.

2. Dugundji actually proved that a linear η can be found such that for each $f \in C^*(A)$, $\eta(f)$ extends f and the range of $\eta(f)$ is contained in the convex hull of the range of f. Furthermore, if one considers bounded, continuous, *complex-valued* functions, then one can obtain complex-linear extenders with the property that the range of $\eta(f)$ is contained in the convex hull of the range of f.

proves Theorem B for certain classes of "generalized metric spaces" which contain mathematically important, but generally non-metrizable, objects such as CW-complexes [C, Bo]. A different kind of extension process was used in [HL_1] to obtain a version of Theorem B for a class of spaces best known for pathological counterexamples, the generalized ordered spaces (see Section 4).

Soon after Dugundji's 1951 paper, Arens [A] presented an example showing that Theorem B cannot be generalized to the class of all normal, or even all compact Hausdorff, spaces and in Section 3 we describe a large class of compact Hausdorff spaces in which one cannot obtain extenders that are both continuous and linear. Recently an example was constructed in [vD] of a regular, first countable, hereditarily Lindelöf space X containing a closed subspace A such that there cannot be a continuous, linear extender from $C^*(A)$ to $C^*(X)$. (This important example disposed of the major open questions about Dugundji extension theory, raised in [M] and [Bo].) Nevertheless, as the positive results in Section 2 show, one can always obtain (i.e., for any closed subspace of a normal space) extenders which are either linear or continuous with respect to the sup-norm topology. (The situation with respect to other function-space topologies is far more confused, as the discussion in Section 4 shows.)

Before proceeding to the positive results in the next section, let us establish certain conventions to be used in this paper. Until otherwise specified (in Section 4), all function spaces carry the sup-norm topology and only bounded functions are considered. The restriction of a function G to a subset A of its domain is denoted by $G \mid A$. All spaces considered in this paper are at least Hausdorff and our topological terminology corresponds to that of [E]. Finally, $\mathbf{R}$ will denote the usual space of real numbers.

The authors wish to thank the referee for his valuable suggestions which substantially improved the organization and exposition of this paper.

2. Positive results. We begin with an easy result whose proof uses only the Axiom of Choice and the Tietze–Urysohn theorem (see [HL_1]).

C. THEOREM. *Let A be a closed subspace of the normal space X. Then there is an extender $\eta: C^*(A) \to C^*(X)$ which is linear.*

Proof. Let B be any basis for the vector space $C^*(A)$. For each $f \in B$ choose an extension $\hat{f} \in C^*(X)$ using the Tietze–Urysohn theorem. Define $\eta(f) = \hat{f}$ and extend η linearly over $C^*(A)$. $\square$

If one is willing to sacrifice linearity of the extension process then one can obtain continuity of the extender η and that $\eta(f)$ will have the same bounds as f, as our next theorem shows.[3] We shall present two proofs of the theorem: one is purely topological and the other is purely analytic.

D. THEOREM. *Let A be a closed subspace of the normal space X. Then there is a continuous extender $\eta: C^*(A) \to C^*(X)$ having the property that $\eta(f)$ has the same bounds as f for each $f \in C^*(A)$.*

Observe that, in order to prove Theorem D, it will be enough to find *any* continuous extender $\hat{\eta}: C^*(A) \to C^*(X)$ since, given any such $\hat{\eta}$, we can obtain the extender η in the theorem by defining $\eta = M \circ m \circ \hat{\eta}$, where M, $m: C^*(X) \to C^*(X)$ are defined by

$$m(f)(x) = \max\{f(x),\ \inf\{f(a) \mid a \in A\}\} \text{ and}$$

$$M(f)(x) = \min\{f(x),\ \sup\{f(a) \mid a \in A\}\}.$$

Clearly M and m are continuous.

3. This theorem solves a problem posed in [HLZ] and [R].

First proof of Theorem D: We begin by reducing the problem to the special case where only compact spaces are considered. If X is not already compact, consider βX, the Čech–Stone compactification of X. Since X is normal, $\bar{A}$, the closure of A in βX, is a topological copy of βA and it is known that each continuous, bounded, real-valued function f on A has a unique extension $\beta f: \beta A \rightarrow \mathbb{R}$ [W, p. 137]. Further, the correspondence $f \rightarrow \beta f$ is an isometry between $C^*(A)$ and $C^*(\bar{A})$. Therefore, if we could obtain a continuous extender η' from $C^*(\bar{A})$ to $C^*(\beta X)$ then we could define $\hat{\eta}: C^*(A) \rightarrow C^*(X)$ by $\hat{\eta}(f) = \eta'(\beta f) | X$. Hence, it is enough to consider the case where X is compact.

Supposing X is compact and A is closed in X, we define $E: A \times C^*(A) \rightarrow \mathbb{R}$ by the rule that $E(a, f) = f(a)$. Then E is continuous [E, p. 139]. Further, since X is compact and $C^*(A)$ is metrizable, $X \times C^*(A)$ is a normal space having $A \times C^*(A)$ as a closed subspace [W, p. 148]. Therefore the Tietze–Urysohn theorem may be applied to find a continuous extension $\hat{E}: X \times C^*(A) \rightarrow \mathbb{R}$ of E (even though E is not bounded; see [W, p. 103]). For each $f \in C^*(A)$, define $\hat{f}$ on X by the rule that $\hat{f}(x) = \hat{E}(x, f)$. Then $\hat{f}$ is continuous and, again using compactness of X, we see that the function η defined by $\hat{\eta}(f) = \hat{f}$ is continuous from $C^*(A)$ to $C^*(X)$. □

Second proof of Theorem D: The Bartle–Graves theorem [BG] asserts that if $r: Y \rightarrow Z$ is a continuous linear surjection, where Y and Z are Banach spaces, then there is a continuous mapping $e: Z \rightarrow Y$ having the property that $e(z) \in r^{-1}(z)$ for each $z \in Z$. To prove Theorem D, let $Y = C^*(X)$, $Z = C^*(A)$ and for each $F \in C^*(X)$ define $r(F) = F | A$. Clearly r is continuous and linear, and the Tietze–Urysohn theorem shows that r is a surjection. The continuous function $e: C^*(A) \rightarrow C^*(X)$ provided by the Bartle–Graves theorem is the required continuous extender. □

E. Corollary. *Let A be a closed subspace of the normal space X. Then $C^*(A)$ is homeomorphic to a closed subspace S of $C^*(X)$ and S is a retract of $C^*(X)$.*

Proof: Let $\eta: C^*(A) \rightarrow C^*(X)$ be the continuous extender found in Theorem D. Any continuous extender is 1-1 and a homeomorphism because the mapping $F \rightarrow (F | A)$, for $F \in \eta(C^*(A))$, is a continuous inverse for η. Finally, the set $S = \eta(C^*(A))$ is closed in $C^*(X)$ since it is a retract of $C^*(X)$, the retraction being the mapping $F \rightarrow \eta(F | A)$ where $F \in C^*(X)$. □

F. Remark: It is possible to generalize Theorem D to the situation where A is an arbitrary subspace of an arbitrary topological space X. For any space X and any subspace $A \subseteq X$, let $E^*(A, X) = \{f \in C^*(A)$: some $f \in C^*(X)$ extends $f\}$. We can show that there is a continuous (bound preserving) extender $\eta: E^*(A, X) \rightarrow C^*(X)$. One proof of the theorem uses the Bartle–Graves theorem as follows: clearly the restriction map $r: C^*(X) \rightarrow C^*(A)$ is a continuous, linear surjection from the Banach space $C^*(X)$ to the normed space $E^*(A, X)$. Therefore, in order to obtain a continuous extender $\eta: E^*(A, X) \rightarrow C^*(X)$ from the Bartle–Graves theorem, all we need to do is to show that $E^*(A, X)$ is a Banach space. That is the point of our next lemma.

G. Lemma. *For any subspace A of X, the space $E^*(A, X)$ is a closed subspace of $C^*(A)$.*

Proof: Suppose $\langle f_n \rangle$ is a sequence in $E^*(A, X)$ which converges to $f \in C^*(A)$ in the sup-norm topology. Then $\langle f_n \rangle$ is a Cauchy sequence and, without loss of generality, we may assume that $\| f_{n+1} - f_n \| < 2^{-n}$ for each $n \geqq 1$. Choose any extension $g_1 \in C^*(X)$ of f_1 having $\| g_1 \| = \| f_1 \|$ and for $n \geqq 2$ choose any extension $g_n \in C^*(X)$ of $(f_n - f_{n-1})$ having $\| g_n \| = \| f_n - f_{n-1} \|$. Since $C^*(X)$ is a Banach space, $g = \Sigma\{g_n : n \geqq 1\}$ belongs to $C^*(X)$ and clearly

$$\| f - g \,|\, A \| \leqq \| f - f_n \| + \Sigma\{\| g_i \| : i \geqq n + 1\} \rightarrow 0$$

so that $f = g \,|\, A$. Hence $f \in E^*(A, X)$. □

REMARK: Corollary E can be generalized, using the result in Remark F above, to assert that for any subspace A of an arbitrary space X, $E^*(A, X)$ is homeomorphic to a closed subset S of $C^*(X)$ where S is a retract of $C^*(X)$.

REMARK: If the space X is paracompact (or even collectionwise normal) then analogues of Theorem D and Corollary E can be obtained for spaces of bounded functions with values in an arbitrary Banach space (instead of $\mathbb{R}$). In addition, analogous results about continuous extension of pseudometrics also fit into the general framework considered in this paper, but the proofs are too technical to present here; see [LP].

3. Examples. In [A] Arens described a compact Hausdorff space X having a closed subset A such that no linear extender $\eta: C^*(A) \to C^*(X)$ can have the property that $\eta(f) \geqq 0$ whenever $f \geqq 0$. In [GS], Gęba and Semadeni proved that if $X = \beta N$ is the Čech–Stone compactification of the discrete space N of natural numbers, and if $A = \beta N - N$, then no linear extender $\eta: C^*(A) \to C^*(X)$ can be continuous. The common feature of these two examples is that, while the compact Hausdorff space X is separable, its closed subspace A contains an uncountable collection of pairwise disjoint relatively open sets. The idea uniting our next two results appears in [GS] and has recently been generalized and exploited by van Douwen [vD]. (See also [HL_2].)

H. THEOREM. *Let X be a separable completely regular space having a closed subspace A which contains an uncountable collection of pairwise disjoint relatively open sets. Then no linear extender from $C^*(A)$ to $C^*(X)$ can be continuous.*

Proof: Suppose some linear extender $\eta: C^*(A) \to C^*(X)$ is continuous. Then there is a natural number n having the property that if $f \in C^*(A)$ then $\|\eta(f)\| \leqq n \cdot \|f\|$. Let $\mathcal{U} = \{U_s \mid s \in S\}$ be an uncountable collection of nonvoid, pairwise disjoint relatively open subsets of A, and for each $s \in S$ choose $a_s \in U_s$. Then choose a function $f_s \in C^*(A)$ having $f_s[A] \subseteq [0,1]$, $f_s(a_s) = 1$, and $f_s[A - U_s] = \{0\}$. Define $V_s = \{x \in X \mid \eta(f_s)(x) > \frac{1}{2}\}$. Each V_s is a nonempty open subset of X so that, X being separable, there is an uncountable $S_0 \subseteq S$ and a point $x_0 \in X$ having $x_0 \in \cap\{V_s : s \in S_0\}$. Choose distinct indices $s_1, \ldots, s_{2n}$ from S_0 and consider the function $g = \Sigma_{i=1}^{2n} f_{s_i}$. Because $U_{s_i} \cap U_{s_j} = \emptyset$ if $i \neq j$, $\|g\| = 1$. Hence

$$n = n \cdot \|g\| \geqq \|\eta(g)\| \geqq \eta(g)(x_0) = \sum_{i=1}^{2n} \eta(f_{s_i})(x_0) > 2n \cdot \tfrac{1}{2} = n$$

and that is impossible. □

I. COROLLARY. *Let X and A be as in Theorem H. Then no linear extender $\eta: C^*(A) \to C^*(X)$ can have the property that $\eta(f) \geqq 0$ whenever $f \geqq 0$.*

Proof: Any such extender would be continuous [KN] and that is impossible by Theorem H. □

The next result sheds light on the extent to which Theorem D could be sharpened and suggests some open questions (see Section 5).

J. THEOREM. *Let X be a separable completely regular space having a closed subspace A containing an uncountable collection of pairwise disjoint relatively open sets. Then no extender $\eta: C^*(A) \to C^*(X)$ can have the property that, for some constant $c < 2$, $\|\eta f - \eta g\| \leqq c \cdot \|f - g\|$. In particular, no extender can be an isometry.*

Proof: Suppose there is a constant $c < 2$ and an extender $\eta: C^*(A) \to C^*(X)$ with the property that $\|\eta f - \eta g\| \leqq c \cdot \|f - g\|$ for each $f, g \in C^*(A)$. Let $\{U_s : s \in S\}$ be an uncountable collection of pairwise disjoint, nonempty, relatively open subsets of A. For each $s \in S$ choose $a_s \in U_s$ and choose two functions $f_s, g_s \in C^*(A)$ such that:

$$f_s[A] \subseteq [0,1], \qquad f_s[A - U_s] = \{0\}, \qquad f_s(a_s) = 1$$
$$g_s[A] \subseteq [-1,0], \qquad g_s[A - U_s] = \{0\}, \qquad g_s(a_s) = -1.$$

Define a set $V_s = \{x \in X : \eta(f_s)(x) > \frac{1}{2}c$ and $\eta(g_s)(x) < -\frac{1}{2}c\}$. Then each V_s is open and $a_s \in V_s$. Since X is separable there is a point p belonging to uncountably many sets V_s. Choose $s, t \in S$ having $p \in V_s \cap V_t$. Then

$$\begin{aligned} c \cdot \| f_s - g_t \| &\geqq \| \eta(f_s) - \eta(g_t) \| \geqq | \eta(f_s)(p) - \eta(g_t)(p) | \\ &= \eta(f_s)(p) - \eta(g_t)(p) > \tfrac{1}{2}c + \tfrac{1}{2}c = c = c \cdot \| f_s - g_t \|, \end{aligned}$$

the last equality holding since $U_s \cap U_t = \varnothing$ forces $\| f_s - g_t \| = 1$. But that is impossible and we conclude that no extender η having $\| \eta(f) - \eta(g) \| \leqq c \cdot \| f - g \|$ for each $f, g \in C^*(A)$ can exist. □

REMARK: In the special case where $c = 1$, Theorem J can also be obtained from Theorem H and the following theorem of Mazur and Ulam [Ba, Theorem 2, p. 166]:

If $E: Y \to Z$ is an isometry from a normed space Y into a normed space Z which sends the zero element of Y to the zero of Z, then E is a linear isometry. (It is easy to see that if an isometric extender $\eta: C^*(A) \to C^*(X)$ exists, then we may obtain an isometric extender which preserves zeroes by defining $E(f) = \eta(f) - \eta(\theta)$, where $\theta \in C^*(A)$ is the zero-function.)

4. Unbounded functions and other function space topologies. For any space Y let $C(Y)$ be the real vector space of all continuous, real-valued, but not necessarily bounded, functions on Y. There are several natural topologies for $C(Y)$, all definable within the following general scheme: let $\mathcal{S}$ be a collection of subsets of Y and suppose $\mathcal{S}$ is closed under finite unions. For $S \in \mathcal{S}$, $\varepsilon > 0$ and $f \in C(Y)$, define

$$N(f, S, \varepsilon) = \{g \in C(Y) : \sup\{|f(y) - g(y)| : y \in S\} < \varepsilon\}.$$

Taking $\mathcal{N}(f) = \{N(f, S, \varepsilon) : S \in \mathcal{S}, \varepsilon > 0\}$ as a neighbourhood base at f, one obtains the topology of *uniform convergence on elements of* $\mathcal{S}$. The three most frequently studied topologies on $C(Y)$ take $\mathcal{S}$ to be, respectively, the singleton $\mathcal{S} = \{Y\}$ (in which case one obtains the *topology of uniform convergence*), the collection $\mathcal{S} = \{K \subset Y : K$ is compact$\}$ (in which case one obtains the *topology of compact convergence* which is equivalent to the *compact-open topology*), and the collection $\mathcal{S} = \{F \subset Y : F$ is finite$\}$ (in which case one obtains the *topology of pointwise convergence*).

In this section we will distinguish these three topologies on $C(Y)$ by using subscripts, namely, $C_u(Y)$, $C_c(Y)$ and $C_p(Y)$ respectively. The set $C^*(Y)$ can be considered as a subspace of any one of these three topological spaces; when $C^*(Y)$ is topologized as a subspace of $C_u(Y)$ we will write $C_u^*(Y)$; the symbols $C_c^*(Y)$ and $C_p^*(Y)$ are defined analogously. Obviously $C_u^*(Y)$ is just the function space $C^*(Y)$ equipped with the topology of the sup-norm which was considered in the first three sections of this paper.

The spaces $C_c(Y)$ and $C_p(Y)$ are topological vector spaces [KN] although not, in general, normable. Unfortunately, $C_u(Y)$ is usually *not* a topological vector space; indeed $C_u(Y)$ is a topological vector space if and only if no continuous real-valued function on Y is unbounded. Clearly, then, functional-analytic tools are more useful in the study of $C_u^*(Y)$ than in the study of $C_u(Y)$ or $C_p(Y)$, for example.

The original Dugundji Extension Theorem for metric spaces provided results for these other vector spaces and for unbounded functions.

K. THEOREM [A, M]. *Let A be a closed subset of a metric space X. Then there is an extender $\eta: C(A) \to C(X)$ such that*:

(1) *η is linear*;

(2) *$\eta(f)$ has the same bounds as f for each $f \in C(A)$;*
(3) *η is continuous if $C(A)$ and $C(X)$ both carry the topology of uniform convergence, the compact-open topology, or the topology of pointwise convergence.*[4]

In Section 2 we showed how, in arbitrary normal spaces, one can obtain linear extenders (Theorem C) or continuous extenders (Theorem D) for spaces of bounded functions with the sup-norm topology. Obviously Theorem C can be generalized to obtain a linear extender from $C(A)$ to $C(X)$ provided A is a closed subspace of a normal space, and our next result is the unbounded analogue of Theorem D.

L. THEOREM. *Let A be a closed subset of a normal space X. Then there is a continuous extender $\eta: C_u(A) \to C_u(X)$ and, under η, $C_u(A)$ is homeomorphic to a closed subspace of $C_u(X)$ which is a retract of $C_u(X)$.*

Proof: The cosets of $C^*(A)$ in $C(A)$ are pairwise disjoint open and closed subspaces of $C_u(A)$; let $S \subseteq C(A)$ have the property that $\{g + C^*(A): g \in S\}$ is exactly the family of distinct cosets of $C^*(A)$ in $C(A)$. Let $\hat{\eta}: C_u^*(A) \to C_u^*(X)$ be the continuous extender found in Theorem D and for each $g \in S$, let $\hat{g} \in C(X)$ be an extension of g (which exists in the light of the Tietze-Urysohn Theorem [E]). Each $h \in C(A)$ is uniquely expressible as $h = g + f$ where $g \in S$ and $f \in C^*(A)$; define $\eta(h) = \hat{g} + \hat{\eta}(f)$. Then η is the required continuous extender. That η is a homeomorphism of $C_u(A)$ onto a retract of $C_u(X)$ follows as in Corollary G and the subsequent Remark. □

When one considers the function spaces $C_c(X)$ and $C_p(X)$, or even their subspaces of bounded functions, the situation becomes far more complex and there is a single space, the Michael line, which seems to provide a counterexample for almost every conjecture. The *Michael line M* is the set of real numbers with the topology $\mu = \{U \cup V: U$ is open in $\mathbf{R}$ and V is any set of irrational numbers$\}$. The space M is first countable and hereditarily paracompact and, while it is not metrizable, one strong version of the Dugundji Extension Theorem can be proved for M, viz., if A is any closed subspace of M then there is a positive, linear, isometric extender $\eta: C_u^*(A) \to C_u^*(M)$.[5] (That fact follows from a theorem in [HL$_1$] since M is an example of a generalized ordered space.) The usual space Q of rational numbers is a closed subspace of M and one can prove:

M. *Example*: With Q and M as above,

1. There is no linear extender $\eta: C(Q) \to C(M)$ such that for each $g \in C(Q)$, $\eta(g)$ has the same bounds (possibly infinite) as g [HL$_1$].

2. There is no continuous extender from $C_c^*(Q)$ to $C_p^*(M)$ or from $C_c(Q)$ to $C_c(M)$ [HLZ].

3. There is no continuous extender from $C_p^*(Q)$ to $C_p^*(M)$ or from $C_p(Q)$ to $C_p(M)$ [HLZ].

The proofs of these results all involve category arguments; the third assertion admits a particularly easy proof so we present it here. (We will actually show that there is no continuous extender from $C_p^*(Q)$ to $C_p(M)$.) Suppose there is a continuous extender $\eta: C_p^*(Q) \to C_p(M)$. As in the Remark following Theorem J, we may assume that $\eta(\theta_Q) = \theta_M$ where θ_Q and θ_M are the zero-functions on Q and M respectively. For each finite subset S of Q and each integer $n \geqq 1$ let $F(S, n) = \{f \in C^*(Q): |f(x)| < 1/n$ for each $x \in S\}$, and let $R(S, n) = \{x \in R:$ if $f \in F(S, n)$ then $|\eta(f)(x)| < 1\}$. Since η is continuous and $\eta(\theta_Q) = \theta_M$, $\mathbf{R} = \cup\{R(S, n): S$ is a finite subset of Q and $n \geqq 1\}$. Since there are only countably many sets $R(S, n)$, and since $\mathbf{R}$ is a complete metric space, there must exist a finite $S_0 \subseteq Q$ and an $n_0 \geqq 1$ such that $\mathrm{Cl}_R(R(S_0, n_0))$ contains an interval (a, b), where $\mathrm{Cl}_{\mathbf{R}}$ denotes closure with respect to the usual space of real numbers. Since S_0 is finite we may choose a point $y \in [(a, b) \cap Q] - S_0$ and then a function $f \in C^*(Q)$ having $f(y) = 2$ while $f(x) = 0$ for each $x \in S_0$.

4. For the topology of uniform convergence, (3) follows from (1) and (2).

5. It is actually proved in [HL$_1$] that one can find a linear extender $\eta: C^*(A) \to C^*(M)$ such that for each $f \in C^*(A)$ the range of $\eta(f)$ is contained in the closed convex hull of the range of f.

Choose a sequence $\langle x_k \rangle$ of points of $R(S_0, n_0)$ which converges to the point y (in the topology of M). Then, since the function $\eta(f)$ is continuous on M, $\langle \eta(f)(x_k) \rangle$ converges to $\eta(f)(y)$ so that, since $|\eta(f)(x_k)| < 1$ for each k, it must be true that $|\eta(f)(y)| \leqq 1$. But that is impossible because $y \in Q$ so that $\eta(f)(y) = f(y) = 2$. $\square$

The situation in which A is a closed subspace of a compact space X is much better if one considers extenders from $C_c(A)$ to $C_c(X)$ since, if X is compact, $C_c(A) = C_u^*(A)$ and $C_c(X) = C_u^*(X)$ showing that Theorem D applies to give continuous extenders. But if one considers the function spaces $C_p(A)$ and $C_p(X)$, the same examples that appeared in Section 3 (namely, compact separable spaces X having closed subspaces A which contain uncountable disjoint collections of relatively open sets) show that one cannot obtain continuous extenders from $C_p(A)$ to $C_p(X)$ [HLZ, Corollary J].

5. Some questions.

(a) Let X be a normal space having a closed subset A. Does there exist an extender $\eta: C^*(A) \to C^*(X)$ such that $\|\eta(f) - \eta(g)\| \leqq 2\|f - g\|$? (Compare Theorem J.)

(b) Let X be a normal space having a closed subspace A. Does there exist a uniformly continuous extender from $C^*(A)$ to $C^*(X)$? (If this question is answered negatively, then so is Question (a).)

Our next question has a very technical formulation but it is part of an easily posed question: which spaces satisfy the version of the Dugundji Extension Theorem given in Theorem B? Obviously any space X for which Theorem B can be proved for *every* closed subspace A must be a normal space; van Douwen [vD] has proved the much stronger theorem that any such space must be hereditarily (collectionwise) normal. The question asks whether the class of spaces satisfying Theorem B is a hereditary class.

(c) Suppose that for each closed subset A of X there is a linear isometric extender $\eta: C^*(A) \to C^*(X)$. Let Y be a subspace of X. Is it true that for each relatively closed $B \subseteq Y$ there is a linear isometric extender from $C^*(B)$ to $C^*(Y)$?

References

[A] R. Arens, Extension of functions on fully normal spaces, Pacific J. Math., 2 (1952) 11–22.

[Ba] S. Banach, Théorie des Opérations Linéaires, Chelsea, New York, 1955.

[BG] R. G. Bartle and L. M. Graves, Mappings between function spaces, Trans. Amer. Math Soc., 72 (1952) 400–413.

[Bo] C. Borges, On stratifiable spaces, Pacific J. Math., 17 (1966) 1–25.

[C] J. Ceder, Some generalizations of metric spaces, Pacific J. Math., 11 (1961) 105–125.

[vD] E. van Douwen, Simultaneous linear extension of continuous functions, General Top. Appl., 5 (1975) 297–319.

[D] J. Dugundji, An extension of Tietze's theorem, Pacific J. Math., 1 (1951) 353–367.

[E] R. Engelking, Outline of General Topology, North Holland, Amsterdam, 1968.

[GS] K. Gęba and Z. Semadeni, Spaces of continuous functions (V), Studia Math., 19 (1960) 303–320.

[KN] J. Kelley and I. Namioka, Linear Topological Spaces, Van Nostrand, Princeton, N.J., 1963.

[HL_1] R. Heath and D. Lutzer, Dugundji extension theorems for linearly ordered spaces, Pacific J. Math., 55 (1974) 419–425.

[HL_2] ——— and ———, The Dugundji extension theorem and collectionwise normality, Bull. Polish Acad. Sci., Ser. Sci. Math. Astron., 22 (1974) 827–829.

[HLZ] ———, ———, and P. Zenor, On continuous extenders, *Studies in Topology*, Academic Press, New York, 1975, pp. 203–214.

[LP] D. Lutzer and T. Przymusiński, Continuous extenders in normal and collectionwise normal spaces, to appear in Fund. Math.

[M] E. Michael, Some extension theorems for continuous functions, Pacific J. Math., 3 (1953) 789–806.

[R] M. E. Rudin, Lectures on set theoretic topology, CBMS Publication Number 23, Amer. Math. Soc., 1975.

[W] S. Willard, General Topology, Addison-Wesley, Reading, Mass., 1970.

MATHEMATICS DEPARTMENT, UNIVERSITY OF PITTSBURGH, PITTSBURGH, PA 15260.
MATHEMATICS DEPARTMENT, TEXAS TECH UNIVERSITY, LUBBOCK, TX 79409.
INSTITUTE OF MATHEMATICS, POLISH ACADEMY OF SCIENCES, WARSAW, POLAND.

DISSECTIONS OF A PLANE OVAL

G. L. ALEXANDERSON AND JOHN E. WETZEL

1. Introduction. A familiar problem in elementary combinatorics is to determine the number R_n of regions formed in a circle by the $\binom{n}{2}$ chords that join n chosen points on the circle, under the assumption that no three of the chords intersect within the circle. Data for $n = 1, 2, 3, 4$, and 5 suggest that $R_n = 2^{n-1}$, but the correct formula is known to be

$$R_n = 1 + \binom{n}{2} + \binom{n}{4}. \tag{1}$$

This paper is concerned with the general plane dissection problem of which this problem is a prototype. In section 3 we establish general formulas for the number R of regions and the number S of segments in a dissected oval in terms of the intersection structure of the dissecting chords. The formula for R is the analog for ovals of a formula for lines in the plane found in 1966 by Alfred Brousseau [**6**]. In section 4 we deduce general formulas for R and S that depend on different data. These formulas are analogs for ovals of formulas for lines in the plane given in 1889 by Samuel Roberts [**15**]. Section 5 concludes with a number of examples, including a third set of general formulas for lines in the plane based on formulas proved in 1826 by J. Steiner [**16**].

In the Klein–Beltrami model of the hyperbolic plane, points are represented by points of the open unit disk and lines by chords of that disk (see, for example, Wylie [**17**, pp. 290–331]). Ovals in the hyperbolic plane are represented by ovals in the disk. It follows that our formulas are correct for dissected ovals in the hyperbolic plane as well as in the Euclidean plane. We shall, however, confine the exposition to the Euclidean plane.

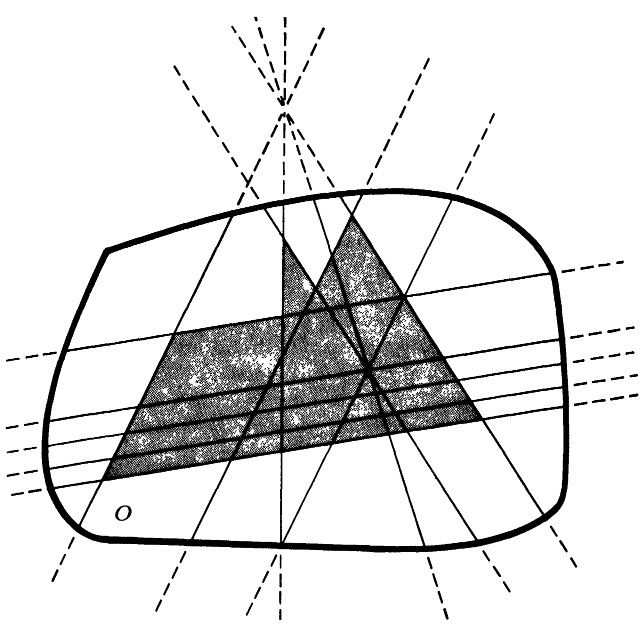


FIG. 1. An oval dissected into 53 regions by a connected arrangement of 12 chords.

2. Dissected ovals. An *oval* is a convex, open, non-empty set in the (Euclidean) plane. Note that an oval need not be bounded. If a line L meets an oval O, the intersection $L \cap O$ is an open line

segment, an open ray, or the full line L; in every case we call $L \cap O$ a dissecting *chord* of O. Suppose that an oval O is cut by an arrangement (i.e., a finite set) α of l chords (Figure 1). To avoid trivial special cases, we assume throughout that $l \geqq 2$, although most of the formulas hold for $l = 1$ as well.

For each point P of O, let $\lambda(P)$ be the number of chords of α that pass through P. If $\lambda(P) \geqq 2$ we say that P is *determined* by α and has *multiplicity* $\lambda(P)$. Let $\mathcal{P}$ be the set of points of O that are determined by α, and suppose that there are $p = |\mathcal{P}|$ such points.

We call a subset of O *bounded in* O (or O-*bounded*) if it is bounded and if its boundary points all lie in O (i.e., if its closure is a compact subset of O); otherwise we call it *unbounded in* O.

The *regions* of α are the connected components of $O - \bigcup(\alpha)$. Each is an open, convex, non-empty subset of O. We denote by R the total number of regions and by R' the number of bounded regions formed by α in O. (The O-bounded regions are shaded in Figure 1.)

The chords of an arrangement α in the oval O are cut by the other chords of α into fragments that may be (open) line segments, (open) rays, or even entire lines. It is convenient to call these the *segments* of α. We denote by S the total number of segments formed by α and by S' the number of these segments that are O-bounded (i.e., open line segments in O both of whose endpoints lie in O).

An arrangement α in an oval O is *connected* if $\bigcup(\alpha)$ is a connected subset of O. The role played by this notion is explained by the following useful result.

LEMMA. *Let α be an arrangement of $l \geqq 2$ chords in an oval O. Then $R - R' \leqq 2l$ and $S - S' \leqq 2l$, and both equalities hold if α is connected.*

Proof. If O is bounded, let $\gamma = \partial O$ be its boundary curve. If O is not bounded, let D be an open disk that is large enough to meet all the chords of α and to contain all of the points in which they intersect; and let $\gamma = \partial(O \cap D)$. Select a reference point X inside γ. There is an $r \in (0, 1)$ sufficiently close to 1 that the homothet Γ of γ with center X and ratio r cuts all the chords of α and surrounds all the points in which they meet (Figure 2). Then the l chords of α meet Γ in $2l$ distinct points, through each of which passes a well-defined unbounded segment of α. The $2l$ distinct points divide Γ into $2l$ distinct subarcs $\Gamma_1, \Gamma_2, \ldots, \Gamma_{2l}$, each of which lies in a well-defined unbounded region of α. Since each unbounded segment and unbounded region has been counted, there are at most $2l$ of each.

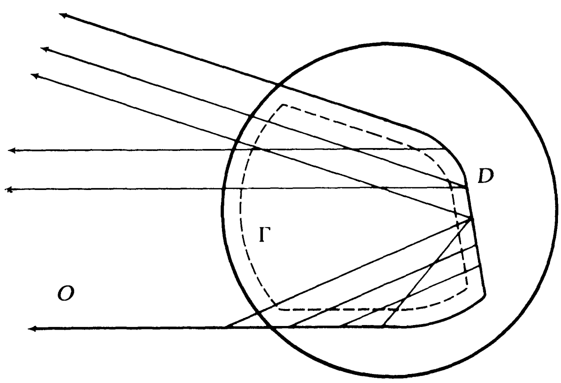


FIG. 2. An unbounded oval dissected by a disconnected arrangement of eight chords.

Now suppose that α is connected. Then every chord is cut by at least one other chord inside γ, and it follows that the $2l$ unbounded segments are distinct. If two different subarcs Γ_i and Γ_j lie in the same region C of α, by convexity there is a line segment in C that joins Γ_i and Γ_j; and adding a segment to each end to reach ∂O (or a ray extending to infinity if O is unbounded in that direction), we have a path that disconnects α. It follows that if α is connected, no two of the subarcs Γ_k lie in the same region of α; and so α has $2l$ unbounded regions.

Conversely, one can see that an arrangement that has $2l$ unbounded regions has to be connected; but there are disconnected arrangements having $2l$ unbounded segments.

3. Brousseau's formulas. Suppose that an oval O is cut by an arrangement α of l chords forming R regions and S segments. Then

$$R = 1 + l + \sum_{P \in \mathscr{P}} (\lambda(P) - 1) \tag{2}$$

$$S = l + \sum_{P \in \mathscr{P}} \lambda(P). \tag{3}$$

Easy proofs can be given by induction on l. Both formulas are trivially correct when $l = 2$. Suppose a new chord L is added to an arrangement α for which the formulas both hold, and suppose that L meets the chords of α in $k \geqq 0$ points in O. The chord L is divided into $k+1$ disjoint segments by the k points of intersection. Since each of these segments divides a previous region into two parts, adding L to α creates exactly $k+1$ new regions and $k+1$ new segments. The right sides of both (2) and (3) plainly increment by exactly $k+1$ when L is taken into account; and this proves both (2) and (3) by induction.

It follows from the lemma that for a connected arrangement,

$$R' = 1 - l + \sum_{P \in \mathscr{P}} (\lambda(P) - 1) \tag{4}$$

$$S' = -l + \sum_{P \in \mathscr{P}} \lambda(P). \tag{5}$$

A different argument for (2) can be given using Euler's formula. Counting the segment-ends of α in two different ways gives (3) at once. Now suppose that O is bounded and that there are $q \leqq 2l$ points of intersection on the boundary ∂O of O. These q points divide ∂O into q subarcs, and the graph formed by the segments of α and these subarcs of ∂O is connected and has data

$$V = q + p = q + \sum_{P \in \mathscr{P}} 1 \qquad \text{vertices,}$$

$$E = l + q + \sum_{P \in \mathscr{P}} \lambda(P) \qquad \text{edges, and} \qquad F = R + 1 \qquad \text{faces.}$$

Now (2) follows from Euler's formula $V - E + F = 2$. If O is not bounded, one can argue similarly on the sphere.

Brousseau [6] gives a very appealing intuitive argument for (2) in the case in which O is the plane, as follows. We argue (3) at the same time. Imagine an auxiliary line B not parallel to any of the lines of α and so situated that all of the points determined by α are on the same side. Then B meets the l lines of α in l distinct points, each of which lies in (and hence counts) a well-defined ray; and these l points divide B into $1 + l$ segments each of which lies in (and hence counts) a well-defined unbounded region of α. Thus the line B initially accounts for l rays and $1 + l$ unbounded regions.

Now sweep the line B parallel to itself across the arrangement. Segments and regions not already counted are encountered by B precisely at the points determined by α, and it is clear that at each point P of multiplicity $\lambda(P)$ exactly $\lambda(P)$ new segments and $\lambda(P) - 1$ new regions are encountered. So in all, $\Sigma_{P \in \mathscr{P}} \lambda(P)$ new segments and $\Sigma_{P \in \mathscr{P}} (\lambda(P) - 1)$ new regions are counted. This proves (2) and (3).

The intuitive idea underlying this argument works equally well for more general ovals, but new segments and regions are also encountered by the sweeping chord at the boundary of the oval. For a very nice example in which the sweep-chord method is used for a convex n-gon, see Freeman [9].

The Euler formula $R-S+p=1$ for an arbitrary dissected oval follows immediately from formulas (2) and (3), and for connected arrangements, the Euler formula $R'-S'+p=1$ follows similarly from formulas (4) and (5).

An important special case occurs when there are no multiple points in the arrangement, i.e., when $\lambda(P)=2$ for each point P of $\mathcal{P}$ (we do not insist that each two chords meet). We call such an arrangement *simple in O*. For simple arrangements, the formulas become

$$R=1+l+p, \qquad S=l+2p \tag{6}$$

and for connected simple arrangements,

$$R'=1-l+p, \qquad S'=-l+2p.$$

In particular, the arrangement in the dissected circle mentioned in section 1 is simple, and there are $l=\binom{n}{2}$ chords and $p=\binom{n}{4}$ points of intersection (because each four points on the circle determine six chords exactly two of which intersect). So (1) follows immediately from (6). (This elegant argument appears in Bauman [**3**], where the first formula of (6) is noted and additional references are cited.)

A closely related problem asks for the number of regions formed in a convex n-gon by its

$$l=\binom{n}{2}-n$$

diagonals if no three of these diagonals meet inside the n-gon. The arrangement is simple and $p=\binom{n}{4}$, so according to (6),

$$R=1+\binom{n}{2}-n+\binom{n}{4}=\binom{n-1}{2}+\binom{n}{4}.$$

A variety of other arguments for this problem are given by Yaglom and Yaglom [**18**, pp. 13, 108–112], Honsberger [**12**, pp. 99–107], and Freeman [**9**]. Some other polygon dissection problems are discussed in [**14**], and a related problem is studied in [**13**].

For n odd, Bol [**5**] and more recently Heineken [**11**] have shown that no three diagonals of the *regular* n-gon are concurrent. It follows that those diagonals partition the interior of the n-gon into

$$\binom{n-1}{2}+\binom{n}{4}$$

regions. For $n\equiv\pm2 \pmod 6$, Harborth [**10**] has found formulas for the number of regions, segments, and points formed in a regular n-gon by its diagonals; but for $n\equiv 0 \pmod 6$ the problem remains open because the intersection structure is not known.

4. Roberts' formulas. Suppose an oval O is cut by an arrangement α of l chords making R regions, S segments, and p points of intersection. Call two chords O-parallel if they are disjoint. For each chord L in α let $\mu(L)$ be the number of chords of α that are O-parallel to L; we call $\mu(L)$ the *order* of L. Let $\mathcal{P}^*=\{P\in\mathcal{P}\mid\lambda(P)>2\}$ be the set of multiple points. Then

$$R=1+l+\binom{l}{2}-\sum_{P\in\mathcal{P}^*}\binom{\lambda(P)-1}{2}-\frac{1}{2}\sum_{L\in\alpha}\mu(L) \tag{7}$$

$$S=l+2\binom{l}{2}-\sum_{P\in\mathcal{P}^*}\lambda(P)(\lambda(P)-2)-\sum_{L\in\alpha}\mu(L) \tag{8}$$

$$p=\binom{l}{2}-\sum_{P\in\mathcal{P}^*}\left[\binom{\lambda(P)}{2}-1\right]-\frac{1}{2}\sum_{L\in\alpha}\mu(L), \tag{9}$$

and if α is connected,

$$R' = 1 - l + \binom{l}{2} - \sum_{P \in \mathcal{P}^*} \binom{\lambda(P)-1}{2} - \frac{1}{2}\sum_{L \in \alpha} \mu(L)$$

$$S' = -l + 2\binom{l}{2} - \sum_{P \in \mathcal{P}^*} \lambda(P)(\lambda(P)-2) - \sum_{L \in \alpha} \mu(L).$$

These formulas can be proved by induction on l, but for variety we deduce them from Brousseau's formulas (2) and (3). The bridge is provided by the nearly obvious identity

$$\binom{l}{2} = \sum_{P \in \mathcal{P}} \binom{\lambda(P)}{2} + \frac{1}{2}\sum_{L \in \alpha} \mu(L), \tag{10}$$

which holds because each two chords in α either do or do not intersect. Indeed, $\binom{l}{2}$ is the number of pairs of chords in α; the sum $\frac{1}{2}\Sigma_{L \in \alpha}\mu(L)$ is the number of pairs of chords in α that do not intersect; and $\Sigma_{P \in \mathcal{P}}\binom{\lambda(P)}{2}$ counts, point by point, the number of pairs of chords in α that intersect.

Now, formula (7) results from rewriting (2) in the form

$$R = 1 + l + \sum_{P \in \mathcal{P}} \binom{\lambda(P)}{2} - \sum_{P \in \mathcal{P}^*} \binom{\lambda(P)-1}{2}$$

and using (10) to eliminate the first sum. Formula (9) is just (10) rewritten as

$$\binom{l}{2} = p + \sum_{P \in \mathcal{P}^*} \left[\binom{\lambda(P)}{2} - 1\right] + \frac{1}{2}\sum_{L \in \alpha} \mu(L)$$

and solved for p; and (8) results from rewriting formula (3) in the form

$$S = l + 2p + \sum_{P \in \mathcal{P}^*} (\lambda(P) - 2)$$

and using (9) to eliminate p.

If O is the entire plane, then two different lines are O-parallel precisely when they are parallel. Suppose in each direction d there are $\pi(d)$ lines of α. If $\pi(d) \geqq 2$ we call d a *multiple direction* of *multiplicity* $\pi(d)$. Let $\mathcal{D}$ be the set of multiple directions. Then

$$\frac{1}{2}\sum_{L \in \alpha} \mu(L) = \sum_{d \in \mathcal{D}} \binom{\pi(d)}{2},$$

since both sums count the total number of non-intersecting pairs of lines in α. So for the plane, formulas (7), (8), and (9) become

$$R = 1 + l + \binom{l}{2} - \sum_{P \in \mathcal{P}^*} \binom{\lambda(P)-1}{2} - \sum_{d \in \mathcal{D}} \binom{\pi(d)}{2} \tag{11}$$

$$S = l + 2\binom{l}{2} - \sum_{P \in \mathcal{P}^*} \lambda(P)(\lambda(P)-2) - 2\sum_{d \in \mathcal{D}} \binom{\pi(d)}{2} \tag{12}$$

$$p = \binom{l}{2} - \sum_{P \in \mathcal{P}^*} \left[\binom{\lambda(P)}{2} - 1\right] - \sum_{d \in \mathcal{D}} \binom{\pi(d)}{2}. \tag{13}$$

These are the formulas Roberts [**15**, pp. 406, 412, 414] gives. Formula (11) was also found independently by A. Blank [**4**, pp. 129, 135–136].

It is an instructive exercise to use Brousseau's sweep-line method to give direct proofs of these three formulas.

Roberts' heuristic arguments for (11) are enlightening. To begin with, l lines in general position divide the plane into $1 + l + \binom{l}{2}$ regions, $2l$ of which are unbounded and

$$1-l+\binom{l}{2}=\binom{l-1}{2}$$

of which are bounded. (These familiar formulas are apparently due to Steiner **[16]**.) An arrangement of lines in the plane can fail to be in general position in two ways: there may be multiple points, or there may be parallels.

Consider first a multiple point P of multiplicity λ. Displace the lines through P a little so as to make an arrangement of λ lines in general position. There are $\binom{\lambda-1}{2}$ bounded regions formed by these λ lines, and all are lost when the lines are brought again to concurrency. So the concurrency of λ lines causes the loss of $\binom{\lambda-1}{2}$ regions.

Now consider a family of μ parallel lines in a direction d. Displace the lines a little so that they are concurrent at a point P far away, making a new multiple point P of multiplicity μ, and let $P\to\infty$ in the direction d. Then $\binom{\mu-1}{2}$ regions are lost at the multiple point P, and $\mu-1$ additional regions are lost beyond P when $P\to\infty$. So the loss due to the μ parallels is

$$\binom{\mu-1}{2}+(\mu-1)=\binom{\mu}{2}.$$

Summing over the multiple points and multiple directions gives Roberts' formula (11). Similar heuristic explanations can be given for (12) and (13).

The reasoning concerning multiple points is essentially local and applies as well to chords in an arbitrary dissected oval, but it is less clear how the heuristics about parallels apply in the more general situation.

5. Examples. In this last section we do a few examples to illustrate the various formulas.

FIGURE 1. First of all, we dispose of the dissected oval pictured in Figure 1. It is divided by a connected arrangement of $l=12$ chords that meet in O to make $p=37$ points, 35 of multiplicity two, one of multiplicity three, and one of multiplicity four. According to Brousseau's formulas,

$$R=1+12+35\cdot 1+1\cdot 2+1\cdot 3=53$$

$$S=12+35\cdot 2+1\cdot 3+1\cdot 4=89$$

$$R'=1-12+35\cdot 1+1\cdot 2+1\cdot 3=29$$

$$S'=-12+35\cdot 2+1\cdot 3+1\cdot 4=65.$$

In this arrangement there are two chords of order two, two chords of order three, seven chords of order four, and one chord of order six. So $\Sigma\mu(L)=44$, and Roberts' formulas give

$$R=1+12+66-(1+3)-\tfrac{1}{2}(44)=53$$

$$S=12+132-(3+8)-44=89$$

$$R'=1-12+66-(1+3)-\tfrac{1}{2}(44)=29$$

$$S'=-12+132-(3+8)-44=65$$

$$p=66-\tfrac{1}{2}(4+10)-\tfrac{1}{2}(44)=37.$$

TRIANGLE. Select x, y, z points on the three (open) sides BC, CA, AB, respectively, of a given triangle ABC and join each to the opposite vertex by a line segment. If no three of these cevians are concurrent inside the triangle, into how many regions is the triangle divided?

The solution is immediate using (6). There evidently are $l=x+y+z$ dividing chords, and they meet in $p=xy+yz+zx$ points inside ABC. So

$$R=1+(x+y+z)+(xy+yz+zx).$$

In the case $x = y = z$, this problem is posed and solved by recursion in Yaglom and Yaglom [**18**, pp. 13, 102]. Analogous results for a dissected tetrahedron in 3-space appear in [**2**] and for a dissected simplex in d-space in [**1**]; see also Zaslavsky [**19**].

COMPLETE REGULAR n-GON. Into how many regions do the extended sides and diagonals of a regular n-gon divide the plane?

Suppose n is odd. Heineken's ingenious argument [**11**] to show that no three diagonals of a regular n-gon are concurrent inside the n-gon shows as well that no three extended sides or diagonals are concurrent outside the n-gon, so the only multiple points in this arrangement are the vertices.

To count the parts it is convenient to use Roberts' formulas, because they do not require prior enumeration of the simple points. Each of the n vertices is a multiple point of multiplicity $n-1$. The $l = \binom{n}{2}$ lines fall into n parallel families each containing $\frac{1}{2}(n-1)$ lines. So, according to formulas (11), (12), and (13),

$$R = 1 + l + \binom{l}{2} - n\binom{n-2}{2} - n\binom{\frac{1}{2}(n-1)}{2} = \frac{1}{8}(n-1)(n^3 - 6n^2 + 21n - 8) \quad \text{regions,}$$

$$S = l + 2\binom{l}{2} - n(n-1)(n-3) - 2n\binom{\frac{1}{2}(n-1)}{2} = \frac{1}{4}n(n-1)(n^2 - 6n + 15) \qquad \text{segments, and}$$

$$p = \binom{l}{2} - \tfrac{1}{2}n^2(n-3) - n\binom{\frac{1}{2}(n-1)}{2} = \frac{1}{8}n(n^3 - 7n^2 + 15n - 1) \qquad \text{points}$$

are formed by the extended sides and diagonals of a regular n-gon when n is odd.

Some results for even n appear in Harborth [**10**].

BURSLEM'S PROBLEM. Suppose an arrangement α of chords in an oval O determines p_k points of each multiplicity $k \geqq 2$, and suppose for each $k \geqq 1$ there are q_k chords each of which meets all but exactly k of the chords of α. Then (10) becomes

$$\binom{l}{2} = \sum_{k=2}^{\infty} \binom{k}{2} p_k + \frac{1}{2}\sum_{k=1}^{\infty} k q_k,$$

which determines l, and Brousseau's or Roberts' formulas give R, S, R', and S'.

The problem of determining R from the data $\{p_k\}$ for an arrangement of lines in the plane without parallels was posed by J. A. Burslem [**7**] and solved (by an application of Euler's formula) by H. Flanders [**8**] (cf. Freeman [**9**]).

STEINER'S FORMULA. Any arrangement α of l lines in the plane falls naturally into $s \geqq 1$ parallel families having $x_1, x_2, \ldots, x_s$ lines respectively, where each $x_i \geqq 1$. In 1826, J. Steiner [**16**] proved by induction that for a simple arrangement α, R is given in terms of the counters $x_1, x_2, \ldots, x_s$ by the formula

$$R = 1 + \sigma_1 + \sigma_2, \tag{14}$$

where σ_1 is the sum of the x_i's and σ_2 is the sum of the $(s/2)$ products $x_i x_j$ with $1 \leqq i < j \leqq s$. This is just formula (6), because there are $l = \sigma_1$ lines, and they clearly meet to form $p = \sigma_2$ points.

In terms of these data, Roberts' formula (11) gives

$$R = 1 + \sigma_1 + \binom{\sigma_1}{2} - \sum_{i=1}^{s} \binom{x_i}{2},$$

and equating this with (14) gives the elementary but not totally trivial identity

$$\binom{x_1 + x_2 + \cdots + x_s}{2} = \sigma_2 + \sum_{i=1}^{s} \binom{x_i}{2}.$$

It is easy to give general formulas for lines in the plane in these terms; one has only to correct for multiple points:

$$R = 1 + \sigma_1 + \sigma_2 - \sum_{P \in \mathcal{P}^*} \binom{\lambda(P) - 1}{2}$$

$$S = \sigma_1 + 2\sigma_2 - \sum_{P \in \mathcal{P}^*} \lambda(P)(\lambda(P) - 2)$$

$$p = \sigma_2 - \sum_{P \in \mathcal{P}^*} \left[\binom{\lambda(P)}{2} - 1\right],$$

and if $s \geqq 2$ so that α is connected,

$$R' = 1 - \sigma_1 + \sigma_2 - \sum_{P \in \mathcal{P}^*} \binom{\lambda(P) - 1}{2}$$

$$S' = -\sigma_1 + 2\sigma_2 - \sum_{P \in \mathcal{P}^*} \lambda(P)(\lambda(P) - 2).$$

The authors would like to thank Marianne Wojciechowski Jankowski for drawing the figures.

References

1. G. L. Alexanderson and John E. Wetzel, Dissections of a simplex, Bull. Amer. Math. Soc., 79 (1973) 170–171.
2. ———, Dissections of a tetrahedron, J. Combinatorial Theory, 11 (1971) 58–66.
3. Norman Bauman, Solution to E 2359, this MONTHLY, 80 (1973) 561–562.
4. Albert A. Blank, Mathematical induction, in Enrichment Mathematics for High School, Twenty-Eighth Yearbook, National Council of Teachers of Mathematics, Washington, D. C., 1963, 118–140.
5. G. Bol, Auflösung von Preisfrage 17, Nieuw Arch. Wisk., (2) 18 (1931) 14–66.
6. Bro. U. Alfred ⟨Brousseau⟩, A mathematician's progress, Math. Teacher, 59 (1966) 722–727.
7. J. A. Burslem, E 1869, this MONTHLY, 73 (1966) 309.
8. Harley Flanders, Solution to E 1869, this MONTHLY, 74 (1967) 864–865.
9. J. W. Freeman, The number of regions determined by a convex polygon, Math. Mag., 49 (1976) 23–25.
10. Heiko Harborth, Diagonalen im regulären n-Eck, Elem. Math., 24 (1969) 104–109.
11. Herman Heineken, Regelmässige Vielecke und ihre Diagonalen, Enseignement Math., (2) 8 (1962) 275–278.
12. Ross Honsberger, Mathematical Gems, Dolciani Mathematical Expositions, No. 1, Mathematical Association of America, Washington, D. C., 1973.
13. Jeanne W. Kerr and John E. Wetzel, Dissections of a triangular prism, Geometriae Dedicata, 4 (1975) 279–289.
14. ———, Dissections of a polygon, to appear in Elem. Math.
15. S. Roberts, On the figures formed by the intercepts of a system of straight lines in a plane, and on analogous relations in space of three dimensions, Proc. London Math. Soc., 19 (1889) 405–422.
16. J. Steiner, Einige Gesetze über die Theilung der Ebene und des Raumes, J. Reine Angew. Math., 1 (1826) 349–364.
17. C. R. Wylie, Jr., Foundations of Geometry, McGraw-Hill, New York, 1964.
18. A. M. Yaglom and I. M. Yaglom, Challenging Mathematical Problems with Elementary Solutions, Vol. 1, Holden-Day, San Francisco, 1964.
19. Thomas Zaslavsky, Maximal dissections of a simplex, J. Combinatorial Theory, 20 (1976) 244–257.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SANTA CLARA, SANTA CLARA, CA 95053.
DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, URBANA, IL 61801.

APPORTIONMENT SCHEMES AND THE QUOTA METHOD

M. L. BALINSKI AND H. P. YOUNG

This paper answers criticisms [4] recently leveled at the Quota Method for Congressional apportionment, and reconsiders the relative merits of various axioms and methods.

1. Background: House-monotonicity and quota. The *apportionment problem* is the problem of determining how to divide the number of representatives in a legislature proportionally among given constituencies. In the United States the problem is rooted in the Constitution, which requires a distribution of Representatives among the various States "*according to their respective numbers.*" The issue is to find an *operational method* for interpreting this mandate, and to identify the essential properties that any fair and reasonable method ought to have. In a recent paper [4] various properties and methods have been suggested as desirable; the purpose of this paper is to examine these proposals in the light of the problem they purport to address.

Formally, the apportionment problem may be stated as follows. Let $\boldsymbol{p} = (p_1, p_2, \ldots, p_s)$ be the populations of *s states*, where each $p_i > 0$ is an integer, and let $h \geqq 0$ be the number of seats in the *house* to be distributed. The problem is to find, for any $\boldsymbol{p}$ and all house sizes $h \geqq 0$, an *apportionment* for h: an s-tuple of nonnegative integers $\boldsymbol{a} = (a_1, \ldots, a_s)$ whose sum is h. A *solution* of the apportionment problem is a function f which to every $\boldsymbol{p}$ and h associates a unique apportionment for h, $a_i = f_i(\boldsymbol{p}, h) \geqq 0$ where $\Sigma_i a_i = h$. A specific apportionment method may give several different solutions, for "*ties*" may occur when using it — for example, when two states have identical populations and must share an odd number of seats. It is useful, for this reason, to define an *apportionment method* $\boldsymbol{M}$ as a non-empty set of solutions. Two different apportionment solutions f and g of a method $\boldsymbol{M}$ may be identical up to some house h and then branch, depending on how a particular tie is resolved. The restriction of f to the domain $(\boldsymbol{p}, h')$, $0 \leqq h' \leqq h$, will be called a *solution up to h*, f^h, and f will be called an *extension* of f^h.

The principles that should apply to apportionment have been intensely debated ever since the Constitutional Convention in 1787. From these debates two basic themes emerge. The first, fundamental to the approaches of Hamilton, Webster, and later contributors is that the ideal or exact number of seats that any state i should receive is $p_i h / \Sigma_j p_j = q_i$, called the *exact quota* of state i, and in any case no state should receive *less* than its *lower quota*, $\lfloor q_i \rfloor$, or *more* than its *upper quota*, $\lceil q_i \rceil$. Any method whose solutions have this property is said to *satisfy quota.*

One such method, first proposed by Alexander Hamilton, and used from 1850 through 1900 is the following: first give to each state i its lower quota and then distribute the remaining seats, one each, to the states with the largest fractional remainders. A fundamental difficulty with this method — which begat the second basic theme for debate — came to light in 1881 when Alabama would have *lost* seats by this method as the house increased from 299 to 300. This behavior is not only shocking to common sense and any reasonable notion of fair division, but has proved to be totally unacceptable politically — as members of Congress immediately perceived. As Representative John C. Bell put it, "*This atrocity which* [mathematicians] *have elected to call a 'paradox' ... this freak presents a mathematical impossibility.*" (Stated in debate, 8 January 1901). For this reason the Hamilton Method was abandoned in 1911, and the basic principle was recognized that an apportionment method must be *house-monotone*; that is, if the total number of seats to be apportioned *increases*, then *ceteris paribus* no state should receive fewer seats than it did before.

Yet in [4] it is said that, "*there is no real reason for requiring apportionment to be house-monotone. The objective should be to minimize inequity.*"

2. Minimizing "inequity" and consistency. Intuitively, "*minimizing inequity*" is what the apportionment problem is all about. The real problem is to determine what "*inequity*" means. To say it is

desirable to "*minimize the length of the inequity vector in Euclidean s-space*" begs the question. Indeed, as pointed out in [**4**], "... *All measures ... of inequity are to some extent arbitrary.*"

Motivated by the need for house-monotone methods E. V. Huntington began in 1921 [**8**] the investigation of several measures of inequity based on pairwise comparisons of states' relative representation. Given populations $\boldsymbol{p}$ and an apportionment $\boldsymbol{a}$ for h, p_i/a_i and a_i/p_i represent, respectively, the *average district size* and the "*share of a representative*" of state i. If $p_i/a_i > p_j/a_j$ then state j is *better off* than state i. Let $T(i,j) = T(j,i) \geqq 0$ be a measure of inequity between states i and j. An apportionment $\boldsymbol{a}$ is *in equilibrium* if no transfer of one seat from a better off state j to a less well off state i reduces the value of $T(i,j)$. Certain T's admit no equilibrium apportionments, but Huntington showed ([**7**] [**8**]) that others do and that five different apportionment methods devolve from these. For example, $T(i,j) = |p_i/a_i - p_j/a_j|$ yields Harmonic Mean (**HM**) apportionments whereas $T(i,j) = |a_j/p_j - a_i/p_i|$ gives Webster (**W**) apportionments. Huntington argued that the most natural choice was the "*relative difference*" $T(i,j) = |p_i/a_i - p_j/a_j|/\min(p_i/a_i, p_j/a_j)$ and showed this choice leads to the Method of Equal Proportions (**EP**). He was persuasive: the U.S. Congress adopted **EP** as the law beginning in 1941. Nevertheless, this choice of measure of inequity remained arbitrary.

Huntington unified his five methods — many of which were anticipated by others in one guise or other — through a computational approach. We generalize it. Let $r(p,a)$ be any real-valued function of two variables called a *rank-index*. Then an apportionment method $\boldsymbol{M}$ is obtained by taking all apportionment solutions f defined recursively as follows:

(i) $f_i(\boldsymbol{p}, 0) = 0,\ 1 \leqq i \leqq s$;

(ii) if $a_i = f_i(\boldsymbol{p}, h)$ and k is some one state for which $r(p_k, a_k) \geqq r(p_i, a_i)$ for $1 \leqq i \leqq s$, then $f_k(\boldsymbol{p}, h+1) = a_k + 1$ and $f_i(\boldsymbol{p}, h+1) = a_i$ for $i \neq k$.

These we call *Huntington Methods.* For example, **HM** has the rank-index $r(p,a) = p/\{2a(a+1)/(2a+1)\}$, $\boldsymbol{W}$ has $r(p,a) = p/(a+\frac{1}{2})$, and **EP** has $r(p,a) = p/\{a(a+1)\}^{\frac{1}{2}}$.

Clearly all Huntington methods are house-monotone. But they also satisfy a condition which epitomizes the very idea of "*method*": namely, the decision as to which state of any pair most deserves the extra seat as the house size is increased by 1 depends only upon the populations and seats already allocated to those states singly, and not on the vector $\boldsymbol{p}$ or the vector $\boldsymbol{a}$ of seats so far allocated. Consider a method $\boldsymbol{M}$ and suppose that it has a solution f allocating to a state with p^* votes a^* seats and to a state with $\bar{p}$ votes $\bar{a}$ seats in a house h, while f allocates to the star state a^*+1 seats and to the bar-state $\bar{a}$ seats in a house $h+1$. Then the star state is said to have *weak priority* at that point and this is written $(p^*, a^*) \gtrsim_{\boldsymbol{M}} (\bar{p}, \bar{a})$. A natural criterion for any method is that the relative claims for an extra seat between two states should depend *only* on their respective populations and apportionments. Specifically, if for some other problem with populations $\boldsymbol{p}'$ there are states having p^* and $\bar{p}$ which are allocated, by a solution of $\boldsymbol{M}$, a^* and $\bar{a}$ seats respectively, and $(\bar{p}, \bar{a}) \lesssim_{\boldsymbol{M}} (p^*, a^*)$, then the states are said to be *tied*, and this is written $(p^*, a^*) \sim_{\boldsymbol{M}} (\bar{p}, \bar{a})$. A method is said to be *consistent* if it treats tied states equally, that is, if $(p^*, a^*) \sim_{\boldsymbol{M}} (\bar{p}, \bar{a})$ implies f^h has both an extension giving the star state a^*+1 seats at $h+1$, and an extension giving the bar state $\bar{a}+1$ seats at $h+1$. Any two states will naturally compare their resultant numbers of seats: a change in priorities could not but be viewed as conflicting with common sense.

THEOREM 1 [**2**]. *An apportionment method $\boldsymbol{M}$ is house-monotone and consistent if and only if it is a Huntington method.*

3. Quotas and pseudo-quotas. It is a major defect of Huntington methods that *none* of them satisfies quota ([**3**], p. 712).

However, one may arrive at certain of Huntington's methods by the device of defining "*pseudo-quotas*" (see [**3**], p. 709). In [**4**], the "*radically different resolution of the Alabama paradox ... apportionment by σ-quota*" ([**4**], last paragraph, p. 684) is one such instance. The approach is to define $p_i/\sigma = q_i(\sigma)$ as the σ-quota of state i, with σ a maximum (*not* minimum as said in [**4**]) allowable average population per district of any state. Letting $a_i(\sigma) = \lceil q_i(\sigma) \rceil$, a σ is sought for which

$\Sigma_1^s a_i(\sigma) = h$ (the size of the house). We omit the improbable case of a tie. Then, $a_i(\sigma) \geqq p_i/\sigma$ or $\sigma \geqq p_i/a_i$. The house-monotone method which results is known as Smallest Divisors (**SD**), was known to Huntington in his 1928 paper [**7**], and was described in precisely this way on p. 709 of [**3**].

The idea of redefining quota in a manner similar to that of the σ-quota is one which is solidly planted in American history. Jefferson advanced it in 1792 (see [**3**], p. 703).

In [**4**] it is suggested that "**EP** *should not be considered unacceptable because it fails to satisfy 'quota' — a short-coming that is easily cured, moreover.*" Presumably this means that an altered quota idea can be used to explain **EP**. This is true. An **EP** apportionment for house size h is found by choosing a σ such that if $a_i(\sigma) = \lfloor \{p_i^2/\sigma^2 + \frac{1}{4}\}^{\frac{1}{2}} + \frac{1}{2} \rfloor$ then $\Sigma_1^s a_i(\sigma) = h$ (see [**3**], p. 709). This hardly seems to commend **EP** — or any method based on some pseudo-quota notion — as a natural method to adopt.

Can satisfying quota be reconciled with monotonicity and consistency? Indeed it can. If consistency is weakened to apply only when upper quota is not violated, then there exists a unique method, the *Quota Method* (**Q**), which satisfies the three properties [**1**], [**3**]. It is said that "**Q** *uses a much more arbitrary and extreme measure of inequity than* **EP**" ([**4**], p. 685). But the fact is that **Q** is not based on *any* measure of inequity. The description of **Q** "*that the augmented representatives shall be as nearly as possible proportional to the populations*" ([**4**], p. 685) is false.* In fact the former defines the Huntington method **J** first proposed by Jefferson (see [**3**], p. 703), also much used in Europe but known as the method of d'Hondt, and cited by Birkhoff as **GD** (Greatest Divisors) which he claims is superior to **Q**.

Q is defined recursively as follows: (i) $f_i(\boldsymbol{p}, 0) = 0, 1 \leqq i \leqq s$; (ii) if $a_i = f_i(\boldsymbol{p}, h)$, $E(h+1)$ is the set of states which can receive an extra seat without violating upper quota at $h+1$, and $k \in E(h+1)$ is some one state satisfying $p_k/(a_k+1) \geqq p_i/(a_i+1)$ for all $i \in E(h+1)$, then $f_k(\boldsymbol{p}, h+1) = a_k + 1$ and $f_i(\boldsymbol{p}, h+1) = a_i$ for $i \neq k$.

4. "Binary fairness". In [**4**] a new apportionment principle called "*binary fairness*" is advanced. This apparently reasonable condition states that if q_i and q_j are the exact quotas of states i and j, and if a_i and a_j are their apportioned numbers of seats, then it should not be possible to transfer a representative from a state i to a state j and reduce $|a_i - q_i| + |a_j - q_j|$. It is, of course, true that Hamilton's method satisfies this condition. But also, we perceive the truth of

THEOREM 2. *An apportionment solution satisfies the binary fairness property if and only if it is a Hamilton method solution.*

COROLLARY. *There exists no house-monotone method satisfying binary fairness.*

This is immediate, since any solution satisfying binary fairness is a Hamilton method solution and no Hamilton method solution is monotone. It can only be concluded that binary fairness is inappropriate to the problem of apportionment.

5. Well-rounding and the Webster method. In [**4**] Birkhoff introduces a condition he calls "binary consistency," and proceeds to attack the quota method **Q** as the only method — of the five proposed by Huntington, the Hamilton method and **Q** — which "fails to have [it]." Here we express this condition in a slightly more natural form and show that it, in fact, uniquely characterizes the Webster method in the class of Huntington methods.

Let $\boldsymbol{a}$ be an apportionment and $\boldsymbol{q}$ the exact quotas. If $a_i > q_i + \frac{1}{2}$ we say that state i's apportionment a_i is *over-rounded*, while if $a_j < q_j - \frac{1}{2}$ that state j's apportionment is *under-rounded.* If there exists no pair of states i and j, with a_i over-rounded and a_j under-rounded, then $\boldsymbol{a}$ is said to be *relatively well-rounded.* This is equivalent to satisfying "binary consistency."

---

* Birkhoff's example is incorrect. If $p_1 = 23{,}500{,}000$ and $p_2 = 1{,}500{,}000$ then **Q** first gives State 2 a second seat when State 1 has 31 seats (not 35).

THEOREM 3. *The Webster method* W *is the unique method that is house-monotone, consistent, and relatively well-rounded.*

Proof. We use the facts (see, e.g., [2]) that: (i) a Webster apportionment a is characterized by

(1) $$\max_i \frac{p_i}{a_i+\frac{1}{2}} \leqq \max_i \frac{p_i}{a_i-\frac{1}{2}} \text{ for } a_i \geqq 1; \text{ and}$$

(ii) Webster apportionments may be found recursively by

(2)
(a) $f(p,0)=0$

(b) if $a=f(p,h)$ and k is some one state for which $p_k/(a_k+\frac{1}{2})=\max_i p_i/(a_i+\frac{1}{2})$ then $f_k(p,h+1)=a_k+1$, $f_i(p,h+1)=a_i$ for $i=k$.

First, that the Webster method is consistent and house monotone is clear by (2). Suppose it is not relatively well-rounded. Then there exists an apportionment a for h, with states i and j satisfying $a_i > q_i+\frac{1}{2}$ and $a_j < q_j - \frac{1}{2}$. Therefore, $a_i - \frac{1}{2} > q_i = p_i h/\Sigma_k p_k$ and $a_j+\frac{1}{2} < q_i = p_j h/\Sigma_k p_k$, implying

$$\frac{p_i}{a_i-\frac{1}{2}} < \frac{\Sigma_k p_k}{h} < \frac{p_j}{a_j+\frac{1}{2}}$$

violating (1). Thus, W satisfies the three conditions.

Conversely, suppose that M is consistent, house-monotone and relatively well-rounded, but is not a set of Webster apportionments. Then there must exist populations p, q having M-apportionments a, b which are Webster apportionments, but

(3) $$(p,a) \gtrsim_M (q,b) \text{ whereas } p/(a+\tfrac{1}{2}) < q/(b+\tfrac{1}{2}),$$

equivalently, $q(a+\frac{1}{2}) \dot{>} p(b+\frac{1}{2})$. By consistency this implies that the two-state problem (p,q) has an M apportionment $(a+1,b)$. But, then, the exact quota of the p-state at $h=a+b+1$ is

$$\frac{p(a+b+1)}{p+q} = \frac{p(a+\frac{1}{2}+b+\frac{1}{2})}{p+q} < \frac{p(a+\frac{1}{2})+q(a+\frac{1}{2})}{p+q} = (a+1)-\tfrac{1}{2},$$

showing that this state is over-rounded. The corresponding exact quota of the q-state is

$$\frac{q(a+b+1)}{p+q} = \frac{q(a+\frac{1}{2}+b+\frac{1}{2}}{p+q} > \frac{p(b+\frac{1}{2})+q(b+\frac{1}{2})}{p+q} = b+\tfrac{1}{2},$$

showing that this state is under-rounded. Therefore, the apportionment $(a+1,b)$ is not relatively well-rounded, a contradiction. This completes the proof.

It should be remembered, however, that in spite of having this property, the Webster method does not satisfy quota.

6. "Bias". The axiomatic approach to apportionment proceeds by making a choice concerning the *principles* which any fair apportionment should satisfy, and then identifying that method (or methods) that satisfy the principles. The advantage of beginning with agreed-upon fairness principles is that subsequent squabbles over particular numbers resulting from these principles are avoided.

Nevertheless, given any method, it is an almost irresistible temptation to analyze particular numerical solutions by adding and substracting different combinations of the numbers to show that the method is in some peculiar sense unfair to certain groups of states. Thus one may question whether a particular solution gives more than a just share to the "*larger*" states versus the "*smaller*" states (or the "*middle*" states) or to the North versus the South, or to the states with large fractions versus those with small fractions, and so forth. These investigations may generally be called ones of "*bias*" and they purport to establish empirically that certain "*new*" principles are violated; principles which by the very nature of the case are different from those already agreed upon as defining the method. For

the notion of bias to even make sense, a normative principle must be postulated; one may then ask what methods (if any) satisfy this principle instead of other principles.

It is stated [**4**] that **SD** is "*unfair to populous states for a simple reason*: *every nonpopulous state 'entitled' to* 1.1 *representatives must be given two representatives* ..." This can only mean that if the exact quota of a "*nonpopulous state*" is at least 1.1 then **SD** assures this state two representatives. This is false, as the following example shows.

						Totals
State populations	4533	4686	5049	6183	9549	30,000
Exact quotas	1.51	1.56	1.68	2.06	3.18	10
SD solution ($\sigma = 4533$)	1	2	2	2	3	10

But, in any case, no normative principle is advanced to support the claim that the numerical example shows bias.

In [**4**] an argument is given "*to show that* **Q** *is biased*" against nonpopulous states. The argument consists of comparing the **H** and **Q** solutions for four 50-state examples, and selecting from the 50 states in each case a subset of "*nonpopulous*" states which **Q** rounds down and **H** rounds up, and a subset of "*populous*" states for which the contrary occurs ([**4**], Tables 2–5).* It is then oberved that **Q** allots less than **H** to the nonpopulous states chosen, and more than **H** to the populous states chosen. It would be as pertinent to remark that virtually any apportionment solution gives some states less than their exact quotas and others more, and that the two sets will in general be different for different methods.

It is true of course that **Q** has a tendency of rounding up the exact quotas of large states more often than those of small states. This is unavoidable — being a necessary consequence of the fairness principles uniquely satisfied by the Quota Method.

If ***H*** is taken as a norm for comparison, ***Q*** is then "*biased*" in that it does not necessarily round up the exact quotas of those states having the largest fractional remainders. This is the procedure which constitutes ***H***, so ***Q*** can hardly but be "*biased*" according to this measure. But ***H*** violates the essential house-monotonicity axiom so cannot be taken as a reasonable norm for comparison. Birkhoff goes on to propose "***H*** *as a good compromise between* **Q** *and* ... **EP**, *which goes so far in the opposite direction that it violates 'quota'* " ([**4**], p. 685). Thus, according to the argument of [**4**], **EP** is also biased, but preferred to ***Q***.

Setting aside the pejorative notion of "*bias*," there *is* a precise sense in which one can talk about one method "*favoring*" large (or small) states ***in comparison with*** another method. This notion is defined in a precise manner and Theorem 1 ([**3**], p. 708) compares the five Huntington methods with respect to "*favoring*" large over small states. But no Huntington method satisfies quota so those comparisons, while interesting, shed little or no light on the supposed "*bias*" of ***Q***.

7. The role of axioms and quota method. The lessons of history clearly point to the necessity of arriving at a fundamental understanding of the *properties* of methods. Put in other terms, political apportionment must be based on principles of fair division rather than on *ad hoc* choices of measures of inequity. Thus axiomatics finds a political role!

In [**4**] Birkhoff attacks ***Q*** for a variety of reasons. First, ***Q*** is faulted because it fails to satisfy the "binary fairness" property, although it is ignored that this property uniquely determines the Hamilton method (which is not house-monotone). Second, ***Q*** is noted to violate "binary consistency," although it is not observed that this property uniquely determines the Webster method (***W***) in the class of Huntington methods (moreover ***W*** is not recommended by Birkhoff). Third, a description of an admittedly arbitrary measure of inequality, supposedly "*used*" by ***Q***, is attacked, but this measure is

* The sets of "*nonpopulous*" or "*populous*" states are different and conveniently chosen in each case.

not used by Q (it characterizes Jefferson's method J, or **GD**, one of three methods recommended by Birkhoff). Fourth, house-monotonicity is discarded as having "*no real reason*," while minimizing any inequity measure is deemed preferable.

This confused state of affairs can only be cleared up through a careful construction of fundamental axioms which satisfy precedents explicitly or implicitly determined by the U.S. Constitution, its framers and interpreters, and by the members of Congress. Further, in the words of Zechariah Chafee, Jr., "*the preservation of a respect for the law will in the long run be best obtained by the adoption of a plan which is least likely to produce a sense of unfairness in those who are forced to obey legislation*" ([**6**], pp. 1043–1044).

This work was supported by the National Science Foundation under Contract MCS 75–07414 A01 with the Graduate School of the City University of New York.

References

1. M. L. Balinski and H. P. Young, A new method for congressional apportionment, Proc. Nat. Acad. Sci. USA, 71 (1974) 4602–4606.

2. ———, On Huntington Methods of Apportionment, to appear as a Research Report, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1976.

3. ———, The quota method of apportionment, this MONTHLY, 82 (1975) 701–730.

4. G. Birkhoff, House-monotone apportionment schemes, Proc. Nat. Acad. Sci. USA, 72 (1976) 684–686.

5. G. A. Bliss, E. W. Brown, L. P. Eisenhart and R. Pearl, Report to the President of the National Academy of Sciences, February 9, Congressional Record, 70th Congress, 2nd Session, 70 (1929) 4966–4967.

6. Z. Chafee, Jr., Congressional Reapportionment, Harvard Law Review, XLII (1929) 1015–1047.

7. E. V. Huntington, The apportionment of representatives in Congress, Trans. Amer. Math. Soc., 30 (1928) 85–110.

8. ———, The mathematical theory of the apportionment of representatives, Proc. Nat. Acad. Sci. USA, 7 (1921) 123–127.

9. M. Morse, J. von Neumann and L. P. Eisenhart, Report to the President of the Academy of Sciences, May 28, 1948.

GRADUATE CENTER, THE CITY UNIVERSITY OF NEW YORK, 33 WEST 42 STREET, NEW YORK, NY 10036, AND THE INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS, LAXENBURG, AUSTRIA.

MATHEMATICAL NOTES

EDITED BY RICHARD A BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Department of Mathematics, University of Wisconsin, Madison, WI 53706.

SOME CHARACTERIZATIONS OF SETS OF MEASURABLE FUNCTIONS

GIULIANA REGOLI

0. Introduction. It is well known that the class of real-valued functions on a set X, which are measurable with respect to a σ-algebra of subsets of X, is an $\mathbf{R}$-algebra of functions which has interesting lattice, algebraic and convergence properties [**3**].

In this paper we shall investigate some of these properties which suffice to characterize sets of measurable functions. Many results in this context are known (for references, see section 5): our purpose is to give reasonably clear, direct and elementary proofs. The most advanced instrument used

is the Weierstrass theorem on uniform approximation of continuous functions by polynomials (on compact intervals of $\mathbb{R}$).

I thank the referee for many helpful suggestions: in particular the statement of the theorem is partly due to him.

1. Background, terminology, notations. Let X be a non-empty set; $\mathbb{R}^X$ is the set of all real-valued functions on X, endowed with the lattice-ordered $\mathbb{R}$-algebra structure given to it by pointwise operations. If $k \in \mathbb{R}$, $\mathbf{k}$ denotes the function on X whose value is constantly k.

If $M \subseteq \mathbb{R}^X$, M^* denotes the subset of M consisting of bounded functions, i.e., those $f \in M$ such that $|f| \leqq \mathbf{k}$ for some $k \in \mathbb{R}$.

The subset M of $\mathbb{R}^X$ is said to have **inversion** when $f \in M$ and $f(x) \neq 0$ for all $x \in X$ implies $1/f \in M$.

A commutative ring A is said to be (von Neumann) **regular** if given $a \in A$ there exists $b \in A$ such that $a = ba^2$. If M is a subring of $\mathbb{R}^X$, then $a = ba^2$ holds if and only if $b(x) = 1/a(x)$ whenever $a(x) \neq 0$. In particular, a regular subring of $\mathbb{R}^X$ has inversion. A **σ-algebra** $\mathscr{A}$ on X is a subset of the power set of X which is closed under complementation and countable union, and such that $X \in \mathscr{A}$.

The set $M(\mathscr{A})$ of $\mathscr{A}$-measurable functions on X consists of all $f \in \mathbb{R}^X$ such that $f^{-1}[U] \in \mathscr{A}$ for every open subset U of $\mathbb{R}$ ($f^{-1}[U]$ denotes the inverse image of U under f).

From [**3**, Th. C, p. 81] it follows that $M(\mathscr{A})$ is a lattice-ordered subalgebra of $\mathbb{R}^X$ which contains $\mathbf{1}$. Moreover, $M(\mathscr{A})$ is closed under pointwise convergence of its sequences ([**3**, Th. A, p. 84]: this is established by proving first that $M(\mathscr{A})$ is closed under pointwise convergence of monotone sequences, and then exploiting the concepts of superior and inferior limit).

A subset S of $\mathbb{R}^X$ is said to be **uniformly bounded** if there exists a constant function $\mathbf{k}$ such that $|f| \leqq \mathbf{k}$ for all $f \in S$.

A lattice subgroup $M \subset \mathbb{R}^X$ is said to be **boundedly σ-closed** (in $\mathbb{R}^X$) if every uniformly bounded countable subset of M has its pointwise supremum (and hence also its infimum) in M.

An argument similar to that sketched above shows that a lattice subgroup $M \subset \mathbb{R}^X$ is boundedly σ-closed if and only if it is closed under pointwise convergence of uniformly bounded sequences; in particular, if M is boundedly σ-closed then it is **uniformly closed,** i.e., M is closed under uniform convergence: if $(f_n)_{n\in\mathbf{N}}$ converges uniformly to $f \in \mathbb{R}^X$ then there exists $n_0 \in \mathbf{N}$ such that the sequence $(f_{n_0+k} - f_{n_0})_{k\in\mathbf{N}}$ is uniformly bounded; clearly this sequence converges (uniformly, hence pointwise) to $f - f_{n_0}$.

If M is boundedly σ-closed in $\mathbb{R}^X$, then M^* is a (conditionally) **σ-complete lattice**, i.e., every countable subset of M with an upper bound in M has a supremum in M; but the converse is in general false (see section 4).

For $f \in \mathbb{R}^X$, write $Z(f) = \{x \in X : f(x) = 0\}$, $Cz(f) = X \backslash Z(f)$; if M is a subset of $\mathbb{R}^X$, then $Z[M] = \{Z(f) : f \in M\}$; analogous meaning for $Cz[M]$. If H is a subset of X, χ_H denotes the characteristic function of H.

Notations are patterned after the book [**1**].

2. Theorem. *Let X be a nonempty set, and let $M \subseteq \mathbb{R}^X$. The following are equivalent:*

(A) *M is the set of all $\mathscr{A}$-measurable functions for some σ-algebra $\mathscr{A}$ on X.*

(B) (1) *M is a vector lattice and $\mathbf{1} \in M$.*

  (2) *M is closed under pointwise convergence of sequences.*

(C) (1) *M is a vector lattice and $\mathbf{1} \in M$.*

  (2) *M is boundedly σ-closed in $\mathbb{R}^X$* (see section 1).

  (3) *M has inversion.*

(D) (1) *M is a regular ring and $\mathbf{1} \in M$.*

  (2) *M is uniformly closed.*

(E) (1) *M is an $\mathbb{R}$-algebra and $\mathbf{1} \in M$.*

  (2) *M is closed under pointwise convergence of sequences.*

Moreover, M uniquely determines the σ-algebra $\mathscr{A}$, which is the set $Cz\,[M]$ ($= Z[M]$) of all cozero-sets of functions in M.

The proof is in section 3. First we need some lemmas. The following lemma makes use of a well-known argument, used in many books to prove the Stone-Weierstrass theorem.

LEMMA 1. *Let M be a uniformly closed $\mathbb{R}$-subalgebra of $\mathbb{R}^X$ containing $\mathbf{1}$. Then M^* is a lattice; if M has inversion, then M is also a lattice.*

Proof. Given $f \in M$, let I be a compact interval of $\mathbb{R}$ containing $f[X]$. By the Weierstrass approximation theorem, there exists a sequence $(p_n)_{n\in\mathbb{N}}$ of polynomials which converges to the function $g(t) = |t|$ uniformly on I. It is clear that $(p_n(f))_{n\in\mathbb{N}}$ is a sequence in M which converges uniformly to $|f|$. Assume that M has inversion: if $f \in M$, we have $f/(1+f^2) \in M^*$, hence $|f/(1+f^2)| \in M^*$, so that $|f| \in M$, because $|f| = |f/(1+f^2)|\,(1+f^2)$.

LEMMA 2. *Let M be a vector sublattice of $\mathbb{R}^X$ containing the constants. Assume that for every $f \in M$ also $\chi_{Cz(f)} \in M$, and that M is uniformly closed. Then $Cz[M] = Z[M] = \mathscr{A}$ is a σ-algebra on X, M^* coincides with the algebra $M^*(\mathscr{A})$ of bounded $\mathscr{A}$-measurable functions, and $M \subseteq M(\mathscr{A})$. If M has inversion, then $M = M(\mathscr{A})$.*

Proof. Given $f \in M$, we have $Cz(f) = Z(\mathbf{1} - \chi_{Cz(f)})$, and $Z(f) = Cz(\mathbf{1} - \chi_{Cz(f)})$, hence $Z[M] = Cz[M] = \mathscr{A}$, and $\mathscr{A}$ is closed under complementation. If $(Cz(f_n))_{n\in\mathbb{N}}$ is a sequence in $\mathscr{A}$, then $\bigcup_{n\in\mathbb{N}} Cz(f_n) = Cz(f)$, where $f = \sum_{n\in\mathbb{N}} |f_n| \wedge \mathbf{2}^{-n}$; clearly this series converges uniformly, hence $f \in M$. This proves that $\mathscr{A}$ is a σ-algebra. Next we show that $M \subseteq M(\mathscr{A})$: in fact if $f \in M$ and $r \in \mathbb{R}$, then $\{x \in X : f(x) > r\} = Cz((f - \mathbf{r}) \vee \mathbf{0}) \in \mathscr{A}$. It follows also that $M^* \subseteq M^*(\mathscr{A})$.

The proof that $M^* \supseteq M^*(\mathscr{A})$ is essentially [3, Th. B, p. 85]: let $f \in M^*(\mathscr{A})$ be given; let $I = [a, b[$, with $a < b$, be an interval containing $f[X]$; for every positive integer n partition I into n pairwise disjoint half-open intervals $I_1, I_2, \ldots, I_n$ of length $(b-a)/n$; put $A_k = f^{-1}[I_k]$. Since f is $\mathscr{A}$-measurable, $A_k \in \mathscr{A}$, i.e., $\chi_{A_k} \in M^*$; hence

$$g_n = \sum_{k=1}^{n} r_k \cdot \chi_{A_k} \in M^*$$

(r_k is the minimum of I_k). It is plain that $|f - g_n| < (b-a)/\mathbf{n}$; by uniform closure, $f \in M$.

Assume now that M has inversion. To show that $M \supseteq M(\mathscr{A})$, we need only to prove that every positive $f \in M(\mathscr{A})$ is in M (recall the decomposition $f = f^+ - f^-$). If $f \geqq \mathbf{0}$, then $\mathbf{1} + f$ is a unit in $M(\mathscr{A})$, hence $g = 1/(\mathbf{1}+f)$ is in $M^*(\mathscr{A}) = M^*$; since g has no zeroes, we have $1/g = \mathbf{1} + f$ in M, hence $f \in M$.

3. Proof of the theorem. That (A) implies (B) and (E) has been remarked in section 1. To show that (A) implies (D), we need only show that $M(\mathscr{A})$ is a regular ring. Given $f \in M(\mathscr{A})$, define $g : X \to \mathbb{R}$ by $g(x) = 1/f(x)$ when $f(x) \neq 0$, and $g(x) = 0$ otherwise. An easy direct proof (or [3, Th. B, p. 81]) shows that $g \in M(\mathscr{A})$. It is obvious that $f = gf^2$. Since a regular ring has inversion we get that (A) implies (C) as well.

(B) implies (A). Given $f \in M$, since $\chi_{Cz(f)}$ is the pointwise limit of the sequence $((n \cdot |f|) \wedge \mathbf{1})_{n\in\mathbb{N}}$, then $\chi_{Cz(f)} \in M$. Evidently M is uniformly closed; by Lemma 2, $M \subseteq M(\mathscr{A})$, and $M^* = M^*(\mathscr{A})$, with $\mathscr{A} = Cz[M]$. If $f \in M(\mathscr{A})$, for $n \in \mathbb{N}$, we have $g_n = (f \wedge \mathbf{n}) \vee (-\mathbf{n})$ in $M^*(\mathscr{A}) = M^*$; since f is the pointwise limit of $(g_n)_{n\in\mathbb{N}}$, we obtain $f \in M$, i.e., $M = M(\mathscr{A})$.

(D) implies (A). We show that M satisfies the hypotheses of Lemma 2. As observed in section 1, a regular subring with $\mathbf{1}$ of $\mathbb{R}^X$ has inversion; it follows that M contains all rational-valued constants, and by uniform closure, all constants, so that M is an $\mathbb{R}$-algebra. By Lemma 1, M is a lattice. Finally, for $f \in M$ there exists $g \in M$ such that $f = gf^2$; it is immediate that $\chi_{Cz(f)} = fg$.

(C) implies (A). Again we show that M satisfies the hypotheses of Lemma 2. First observe that for

f in M, $\chi_{Cz(f)}$ is the pointwise supremum of the set $\{(n \cdot |f|) \wedge 1 : n \in \mathbf{N}\}$, which is uniformly bounded by 1. As observed in section 1, every boundedly σ-closed group is uniformly closed.

(E) implies (B) (hence (E) implies (A)). We only need to prove that M is a lattice. Given $f \in M$, we prove that $|f| \in M$. By the Weierstrass approximation theorem, for every positive integer n we have a polynomial p_n such that

$$\left| |t| - p_n(t) \right| \leqq 2^{-n}$$

for every $t \in [-n, n]$. It is immediate that $p_n(f)_{n \in \mathbf{N}}$ is a sequence in M which converges pointwise to $|f|$.

4. Remarks. If a topology on X is given, denote by $C(X)$ the set of all real-valued continuous functions. It is well known that $C(X)$ is a uniformly closed lattice-ordered $\mathbf{R}$-algebra with inversion. From [**1**, problem 4, J] it follows that $C(X)$ is a regular ring if and only if X is a P-space (i.e., a space in which every G_δ is open). Thus condition (D) of the theorem shows that $C(X)$ is an algebra of measurable functions if and only if X is a P-space.

From [**1**, problem 3, N] it follows that $C(X)$ is a (conditionally) σ-complete lattice if and only if X is basically disconnected, i.e., the closure of every cozero-set of X is open. There exist many examples of basically disconnected (even extremally disconnected, see [**1**, problem 1, H]) spaces which are not P-spaces, the simplest one being perhaps the space Σ of [**1**, problem 4, N].

This shows that hypothesis (C(2)) of the theorem cannot be replaced by σ-completeness.

Let V be a vector sublattice of $\mathbf{R}^X$ with $\mathbf{1}$ in V. Let $\mathscr{B}$ be the σ-algebra on X generated by $Cz[V]$ (equivalently, by $Z[V]$). Then $M(\mathscr{B})$ is a vector lattice which contains V and is closed under pointwise convergence of sequences. If M is such a vector sublattice of $\mathbf{R}^X$, then, by the Theorem, $M = M(\mathscr{A})$ where $\mathscr{A} = Cz[M] \supseteq Cz[V]$; it follows that $\mathscr{A} \supseteq \mathscr{B}$, i.e., $M(\mathscr{A}) \supseteq M(\mathscr{B})$. That is to say, $M(\mathscr{B})$ is the smallest vector sublattice of $\mathbf{R}^X$ which contains V and is closed under pointwise convergence of sequences. We could call $M(\mathscr{B})$ the set of V-Baire functions on X, $B(V)$, (see [**5**, 12. H]). Starting with $V = C(X)$, then $B(V)$ is precisely the set of Baire functions on X, i.e., functions which are measurable with respect to the σ-algebra generated by the zero-set of the topological space X (compare with [**4**, 11.46]).

5. References. The theorem we prove (section 2) belongs to a category of theorems which could be considered more or less known.

The equivalence of (A) and (B) can be found in [**5**, 12.H] (but here the focus is on the Daniell integral, measurability is only a by-product).

Hager in [**2**] proves, among other things, that (A) and (D) are equivalent: his setting is much more abstract than ours, and requires representation theorems for ϕ-algebras. Once we accept the fact that a real-semisimple ϕ-algebra is an algebra of real-valued functions, then the equivalence of (a) and (c) of Theorem 2.3 in [**2**] also follows from our theorem. But we don't get condition (b) of [**2**, Th.3.3].

References

1. L. Gillman and M. Jerison, Rings of Continuous Functions, Van Nostrand, Princeton, N. J., 1960.
2. A. W. Hager, Algebras of measurable functions, Duke Math. J., 38 (1971) 21–27.
3. P. R. Halmos, Measure Theory, Van Nostrand, Princeton, N.J., 1950.
4. E. Hewitt and K. Stromberg, Real and Abstract Analysis, Springer-Verlag, New York, 1965.
5. L. H. Loomis, An Introduction to Abstract Harmonic Analysis, Van Nostrand, Princeton, N.J., 1953.

Istituto di Matematica, Via Vanvitelli 1, 06100 Perugia, Italy.

WHITLEY'S TECHNIQUE AND K_δ-SUBSPACES OF BANACH SPACES

G. J. O. JAMESON

Introduction. By a "K_δ-subspace" of a Banach space we mean a subspace that is the intersection of the kernels of countably many continuous linear functionals. It is elementary that either of the following statements is equivalent to E being a K_δ-subspace of X: (i) E is the kernel of a continuous linear mapping from X to m, or (ii) the space X/E has a countable, separating family of continuous linear functionals.

In [5], R. J. Whitley gave an elegant proof of Phillips' theorem that c_0 is not complemented in m. What Whitley really showed is that c_0 is not a K_δ-subspace of m. We show here that a rather stronger statement about K_δ-subspaces is implied by Whitley's reasoning. This is our Proposition 1.

One application of our result is a short, measure-free proof of the main result of Rosenthal [4] (for countable Γ). (It should be mentioned that the underlying measure-theoretic lemma of [4] has been greatly simplified by Kupka [2].) Rosenthal's theorem can be used, in turn, to derive a more general version of Proposition 1, involving *copies* of c_0 and m instead of the spaces themselves.

Notation. The set of positive integers is denoted by $\mathbf{N}$, and the sequence spaces m, c_0 have their usual meaning (the scalars may be real or complex). We regard numerical sequences as functions on $\mathbf{N}$, writing $x(n)$ for the "nth term" of x. For $N' \subseteq \mathbf{N}$, we write $m(N')$ for the set of $x \in m$ having $x(i) = 0$ for $i \in \mathbf{N}\backslash N'$. The symbol e_k denotes the sequence having 1 in place k and 0 elsewhere. The dual of a Banach space X is denoted by X^*.

PROPOSITION 1. *Suppose that E is a K_δ-subspace of m that contains c_0. Then E contains $m(N_1)$ for some infinite set $N_1 \subseteq \mathbf{N}$.*

Proof. There exists a collection $\{N_\lambda : \lambda \in \mathbf{R}\}$ of infinite subsets of $\mathbf{N}$, indexed by the reals, such that the intersection of any two is finite. For f in m^*, define $Q(f)$ to be the set of λ such that there exists x_λ in $m(N_\lambda)$ with $f(x_\lambda) \neq 0$. We shall show that if f vanishes on c_0, then $Q(f)$ is countable. If E is the inersection of the kernels of functionals $f_n (n \in \mathbf{N})$, it will follow that there is a λ not in any $Q(f_n)$. This means that each f_n vanishes on $m(N_\lambda)$, so that E contains $m(N_\lambda)$. (Of course, this actually holds for uncountably many different λ.)

Suppose, then, that $\|f\| = 1$ and f vanishes on c_0. Choose $k \in \mathbf{N}$, and let $\lambda_1, \ldots, \lambda_s$ be distinct real numbers such that, for $1 \leqq i \leqq s$, there is an element x_i of $m(N_{\lambda_i})$ with $\|x_i\| = 1$ and $f(x_i) \geqq 1/k$. We shall prove that $s \leqq k$, from which it will follow that $Q(f)$ is countable, as stated. Define an element y of m as follows. If j belongs to exactly one (say N_{λ_p}) of $N_{\lambda_1}, \ldots, N_{\lambda_s}$, let $y(j) = x_p(j)$. Let $y(j) = 0$ for all other j. Then $y(j) = (x_1 + \cdots + x_s)(j)$ unless j is in one of the sets $N_{\lambda_p} \cap N_{\lambda_q}$. Since there are only finitely many such j, the element $y - (x_1 + \cdots + x_s)$ belongs to c_0, so that

$$f(y) = f(x_1 + \cdots + x_s) \geqq s/k.$$

But $\|y\| \leqq 1$, so we must have $s \leqq k$, as stated.

In other words, if (f_n) is a sequence of functionals on m, each taking the value 0 on c_0, then there is an infinite set N_1 such that each f_n takes the value 0 on $m(N_1)$. The statement is of some interest even for one functional.

COROLLARY 1.1. *If G is a separable subset of m^* and each member of G takes the value 0 on c_0, then there is an infinite set N_1 such that each member of G takes the value 0 on $m(N_1)$.*

Proof. Some countable set $\{g_n\}$ is dense in G. If each g_n takes the value 0 on $m(N_1)$, then, clearly, so does each member of G.

Let f be an element of m^*. By the well-known characterization of c_0^*, the series $\sum_{k=1}^{\infty} |f(e_k)|$ is

convergent to $\|f|_{c_0}\|$, and we have

$$f(x) = \sum_{k=1}^{\infty} x(k) f(e_k)$$

for all $x \in c_0$. We can use the same formula to define a functional $\tilde{f}$ on m: let $\tilde{f}(x) = \sum_{k=1}^{\infty} x(k) f(e_k)$ for $x \in m$. Then $\tilde{f}$ is the natural extension to m of $f|_{c_0}$; in measure-theoretic terms, it is the "atomic part" of f. The mapping $f \to \tilde{f}$ is obviously linear and norm-reducing. With this notation, we have:

COROLLARY 1.2. *Let G be a separable subset of m^*. Then there is an infinite set $N_1 \subseteq \mathbf{N}$ such that $\tilde{g}(x) = g(x)$ for all $g \in G$ and $x \in m(N_1)$.*

Proof. Apply Corollary 1.1 to the separable set $\{g - \tilde{g} : g \in G\}$.

This corollary is quite similar to a result (31.2(2)) used in **[1]** as a stage in the proof of the theorem known as Phillips' Lemma. Using our corollary, one can complete the proof of Phillips' Lemma in roughly similar fashion.

We now show how Rosenthal's result (**[4]**, Proposition 1.2) can be derived from Proposition 1:

PROPOSITION 2. *Let Y be a normed linear space, and let T be a continuous linear mapping of m into Y such that $T|_{c_0} (= T_0)$ is an isomorphism. Then there is an infinite set $N_1 \subseteq \mathbf{N}$ such that $T|_{m(N_1)} (= T_1)$ is an isomorphism and $\|T_1^{-1}\| \leqq \|T_0^{-1}\|$.*

Proof. Let δ_n be the element of m^* defined by: $\delta_n(x) = x(n)$. Define a linear functional f_n on $T(c_0)$ by putting

$$f_n(Tx) = \delta_n(x). \tag{1}$$

Then

$$|f_n(Tx)| \leqq \|x\| \leqq \|T_0^{-1}\| \cdot \|Tx\|,$$

so $\|f_n\| \leqq \|T_0^{-1}\|$. Extend f_n to an element of Y^* (still denoted by f_n) with the same norm. Then (1) says that each of the functionals $T^*f_n - \delta_n$ vanishes on c_0. By Proposition 1, there is an infinite set N_1 such that each of these functionals vanishes on $m(N_1)$; in other words, (1) holds for all $x \in m(N_1)$. So for $x \in m(N_1)$ we have

$$\begin{aligned}\|x\| &= \sup_n |\delta_n(x)| \\ &= \sup_n |f_n(Tx)| \leqq \|T_0^{-1}\| \cdot \|Tx\|.\end{aligned}$$

This proves the result.

Using this, we now prove the promised more general version of Proposition 1 (note, however, that Proposition 1 is not a special case). "Copy" means "isomorphic copy."

PROPOSITION 1′. *Let X be a Banach space that is complemented in its second dual, and let E be a K_δ-subspace of X that contains a copy of c_0. Then E contains a copy of m.*

Proof. Let $E = \bigcap_{n=1}^{\infty} \ker f_n$, and let V_0 be an isomorphism of c_0 onto a subspace of E. Then there is a continuous linear mapping V, extending V_0, of m into X (let $V = PV_0^{**}$, where P is a projection of X^{**} onto X). Each functional $V^*(f_n)$ vanishes on c_0, so, by Proposition 1, there is an infinite set N_1 such that each of these functionals vanishes on $m(N_1)$. This means that V maps $m(N_1)$ into E. By Proposition 2, applied to $m(N_1)$, there is an infinite subset N_2 of N_1 such that V is an isomorphism on $M(N_2)$. Hence E contains a copy of m.

Compare Corollary 1.5 of **[4]**, where the same is asserted for *complemented* subspaces of X. Any

dual space is complemented in its own second dual, so Proposition 1′ generalizes the following well-known result of Bessaga and Pełczyński: if a dual space contains a copy of c_0, then it contains a copy of m. Note that we can deduce that $C(I)$ (for example) is not complemented in its second dual.

For completeness, we repeat from [4] the analogous generalization of Proposition 2 ([4], Theorem 1.3):

PROPOSITION 2′. *Let X be a Banach space that is complemented in its second dual, and let Y be a normed linear space. Suppose that T is a continuous linear mapping from X to Y, and that there is a copy E of c_0 such that $T|_E$ is an isomorphism. Then there is a copy F of m such that $T|_F$ is an isomorphism.*

Proof. Let V_0 be an isomorphism of c_0 onto E. As before, there is a continuous linear mapping V, extending V_0, of m into X. Then $TV|_{c_0}$ is an isomorphism, so by Proposition 2, there is a copy A of m such that $TV|_A$ is an isomorphism. Let $F = V(A)$.

Either of these two propositions provides an alternative to the work in [3]. For once it is known that an infinite-dimensional complemented subspace of m must contain a copy of c_0, either proposition tells us that it must contain a copy of m; this is what is proved in [3].

Several further applications of Proposition 2′ are given in [4].

Problem. Describe the infinite-dimensional K_δ-subspaces of m. In particular, are there any separable ones?

References

1. G. Köthe, Topologische lineare Räume I, Springer, 1960.

2. J. Kupka, A short proof and generalization of a measure theoretic disjointization lemma, Proc. Amer. Math. Soc. 45 (1974) 70–72.

3. J. Lindenstrauss, On complemented subspaces of m, Israel J. Math., 5 (1967) 153–156.

4. H. P. Rosenthal, On relatively disjoint families of measures, with some applications to Banach space theory, Studia Math., 37 (1970) 13–36.

5. R. J. Whitley, Projecting m onto c_0, this MONTHLY, 73 (1966) 285–286.

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY, LANCASTER LA1 4YL, GREAT BRITAIN.

A GRAPHIC METHOD FOR CONSTRUCTING SPHERICAL INDICATRICES

LAWRENCE L. LEE

Abstract. A graphic method is presented for constructing the spherical indicatrices of the tangent, principal-normal, and binormal of space curves described by intrinsic equations. Using this method, it will not be necessary to integrate the Serret–Frenet formulas. The indicatrices are obtained directly by mapping plane curves onto a unit sphere. This mapping requires the sphere to roll on the plane such that the locus of the contact point in the plane is the preimage and the corresponding locus on the sphere is the desired indicatrix. Equations for the plane curves are derived and procedures for performing the graphic mapping are outlined.

Introduction. The locus of a point whose position vector is the tangent to a curve $\mathscr{C}$ is called the spherical indicatrix of the tangent to $\mathscr{C}$. The spherical indicatrices of the principal-normal and the binormal to $\mathscr{C}$ are similarly defined, (see [1]).

The tangent, principal-normal, and binormal can be obtained from the intrinsic equations by integration of the Serret–Frenet formulas. Then their respective indicatrices can be constructed. The indicatrices can also be constructed graphically, without need for integrating the Serret–Frenet formulas. This graphic construction makes use of the fact that the indicatrices are spherical curves; so

they can be constructed by "rolling" plane curves onto a sphere. This is generally easier than plotting the space curves directly. Furthermore, the equations for the plane curves are also simpler than the Serret–Frenet formulas.

The construction of the principal-normal indicatrix will be described first, then it will be shown that similar constructions give the tangent and binormal indicatrices.

The "Roll" condition. If a spherical curve $\mathcal{R}(t)$ with a frame field $e_i(t)$ $(i = 1,2)$ along it is rolled without pivoting or slipping onto a plane curve $\mathcal{R}'(t)$ with a frame field $e_i'(t)$, then

$$\frac{de_i}{dt}\cdot e_j = \frac{de_i'}{dt}\cdot e_j'. \tag{1}$$

Application to the principal-normal indicatrix. The principal-normal indicatrix $\mathcal{N}(s)$ of a space curve $\mathcal{C}(s)$ is a spherical curve (s is the arc-length along $\mathcal{C}$). The frame field along $\mathcal{N}(s)$ are $t(s)$ and $b(s)$. They satisfy the Serret–Frenet formulas

$$dt/ds = \kappa n \tag{2}$$

and

$$db/ds = -\tau n. \tag{3}$$

When $\mathcal{N}(s)$ is rolled onto the plane curve $\mathcal{N}'(s)$, $n(s)$ becomes normal to the plane. Equations 1, 2 and 3 give the corresponding equations for the frame field t' and b' along $\mathcal{N}'$:

$$\frac{dt'}{ds} = \frac{db'}{ds} = 0 \tag{4}$$

which suggest that t' and b' are suitable for use as basis for a cartesian coordinate system. In this coordinate system, the equation for $\mathcal{N}'$ is derived from the Serret–Frenet formula for n: The parametric equations for $\mathcal{N}'$ are

$$dx = -\kappa\, ds, \tag{5}$$

$$dy = \quad \tau\, ds. \tag{6}$$

They enable us to construct $\mathcal{N}'$ directly from the intrinsic equations for $\mathcal{C}$. The principal-normal indicatrix is obtained by rolling $\mathcal{N}'$ onto a unit sphere as described below.

Graphic construction of the principal-normal indicatrix. The graphic construction is a straightforward implementation of "letting $\mathcal{N}'$ roll onto $\mathcal{N}$". Physical realizations of the geometric objects are used. This does not impose any limit on the achievable accuracy of the result. To produce accurate results, the plane curve should be plotted on thin, strong, and stretch-resistant material (certain foils may be better than paper for this purpose). A precision sphere should be used. The sphere is not consumed, and it can be reused for other constructions of this type.

Let $\mathcal{N}'$ be plotted, and a neighborhood of $\mathcal{N}'$ (the union of all sufficiently-small neighborhoods of points on $\mathcal{N}'$) be isolated by cutting out a strip of the paper around $\mathcal{N}'$. This cut strip should ideally be only infinitesimal in width, but a wider strip must be used in order to maintain rigidity within the plane. A certain amount of "art" is needed for cutting a strip of the optimum width. To assure that no slipping or pivoting occurs, and that contacts are indeed made at points on $\mathcal{N}'$, a thin line of contact adhesive may be applied along $\mathcal{N}'$, and the strip of paper allowed to become glued onto the sphere as it rolls along. Care should also be taken to assure that the strip conforms to the spherical surface without wrinkling.

As the strip of paper containing $\mathcal{N}'$ is transferred into the tangent space of the sphere, $\mathcal{N}'$ is mapped into the principal-normal indicatrix. If the vectors t' and b' had been plotted along $\mathcal{N}'$, they would have become t and b respectively.

The tangent and binormal indicatrices. The tangent indicatrix $\mathcal{T}(s)$ and the binormal indicatrix $\mathcal{B}(s)$ are spherical curves and can also be rolled onto plane curves that can be constructed directly from the intrinsic equations of $\mathcal{C}$. Let $n(s)$ and $b(s)$ be the frame field along $\mathcal{T}(s)$ which is rolled onto the plane curve $\mathcal{T}''(s)$ with frame field $n''(s)$ and $b''(s)$. Then the Serret–Frenet formulas and equation 1 give

$$\frac{dn''}{ds} = \tau b'' \tag{7}$$

$$\frac{db''}{ds} = -\tau n'' \tag{8}$$

which shows that this frame field rotates at a rate τ. But the tangent of $\mathcal{T}''$ is in the direction n'', and the arc-length (s'') along $\mathcal{T}''$ is related to the arc-length (s) along $\mathcal{C}$ by $ds'' = \kappa\, ds$, so the curvature of $\mathcal{T}''$ is τ/κ. This describes $\mathcal{T}''$ in intrinsic form.

Similar analyses give the intrinsic equation for $\mathcal{B}'''$, the image of the binormal indicatrix $\mathcal{B}$ rolled onto a plane: The element of arc-length along $\mathcal{B}'''$ is $ds''' = |\tau|\, ds$, and the curvature of $\mathcal{B}'''$ is κ/τ.

Since κ is non-negative, $\mathcal{N}'$ cannot enclose any area. This is very convenient for the graphic mapping of the principal-normal indicatrix. But there is no guarantee that $\mathcal{T}''$ and $\mathcal{B}'''$ will not be self-crossing. If one of these curves crosses itself, it must be made to cross *over* or *under* itself; otherwise part of the curve would enclose an area in the plane, then it cannot be mapped onto the sphere because it cannot satisfy the Gauss–Bonnet theorem. The curve can be made to cross over or under itself by constructing its physical realization in more than one non-crossing section and have these sections joined together at positions between the crossings *after* cutting out the strip.

Reference

1. T. J. Willmore, An Introduction to Differential Geometry, Clarendon Press, Oxford, 1966, p. 29.

THE MAGNAVOX COMPANY, FORT WAYNE, IN 46804.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics and Statistics, The University of Calgary, Calgary, Alberta, Canada, T2N 1N4.

MORTALITY OF 2×2 MATRICES

P. SCHULTZ

Let H be a non-empty finite set of $n \times n$ non-zero complex matrices. H is called *mortal* if there exists a product $A_1A_2\cdots A_k$ of elements of H equal to the zero matrix. The *decision problem for mortality* is to find an algorithm which, given H, decides whether H is mortal. For example, if $n = 1$, H is never mortal, so a trivial algorithm exists.

By an ingenious construction, Michael S. Paterson [4] showed that in the case $n = 3$, the decision problem for mortality is equivalent to Post's Correspondence Problem, so is unsolvable. His method works for all $n \geqq 3$, so the only unsolved case is $n = 2$, the problem of the title.

It is easy to see that H is mortal if and only if some product $A_1 A_2 \cdots A_k = 0$ with A_1 and A_k of rank 1, and the rest of rank 2, so the problem is equivalent to the following:

PROBLEM: *Find an algorithm which, given a finite set H of non-singular linear transformations of the complex plane, and lines L and M through the origin, determines whether some product from H maps L onto M.*

Since $GL(C, 2)$ is 2-fold transitive on lines, there are infinitely many choices for a matrix mapping L onto M, so this problem is probably too hard to attack directly; however, some reductions are possible. Firstly, multiplying any matrix by a non-zero scalar does not change its effect on lines, so the problem is equivalent to the following:

Find an algorithm which, given a finite set of generators of a subsemigroup S of PSL(C, 2), (considered as acting on the complex projective line P′), and points L, M in P′, decides whether M is in the orbit of L under S.

What has been gained by this reduction is that it is known ([1], [2] and [3]) that certain finitely generated subgroups of $PSL(C, 2)$ are free products of finite cyclic groups; for example, ([3], p. 139) $PSL(Z, 2)$ is the free product $[x] * [y]$ of a cyclic group $[x]$ of order 2 and a cyclic group $[y]$ of order 3, where x is the element of $PSL(Z, 2)$ which corresponds to the matrix

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

and y is the element of $PSL(Z, 2)$ which corresponds to the matrix

$$\begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}.$$

This means that every element A of $PSL(Z, 2)$ has a unique decomposition of the form:

$$A = x^a y^{b_1} x y^{b_2} x \cdots x y^{b_n},$$

where $a = 0$ or 1, $b_i = 1$ or 2 for $1 \leqq i \leqq n - 1$, and $b_n = 0, 1$ or 2. It is not known, however, whether all finitely generated subgroups of $PSL(C, 2)$ are free products of cyclic groups.

The decision problem can be solved in some special cases, for example when H contains only upper triangular matrices; a positive solution in general would be of interest to both geometers and algebraists. A negative solution would provide a new type of unsolvable problem.

References

1. J. Lehner and M. Newman, Real two-dimensional representations of the modular group and related groups, Amer. J. Math., 87 (1965) 945–954.

2. ——— and ———, Real two-dimensional representations of the free product of two finite cyclic groups, Proc. Cambridge Philos. Soc., 62 (1966) 135–141.

3. M. Newman, Integral Matrices, Academic Press, New York, 1972, p. 139.

4. M. S. Paterson, Unsolvability in 3 × 3 matrices, Studies in Applied Mathematics, 49 (1970) 105–107.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WESTERN AUSTRALIA, NEDLANDS, 6009, AUSTRALIA.

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Department of Mathematics, University of Wisconsin, Madison, WI 53706.

FORMULAE FOR THE ARC-LENGTH OF A CURVE IN R^N

M. J. PELLING

Let G denote a parametrized continuous curve $(f_1(t),\dots,f_n(t))$ in R^n with parameter t ranging in an interval $[0,c]$. For rectifiable G the formula

$$\text{(A)} \qquad s=\int \sqrt{f_1'(t)^2+\cdots+f_n'(t)^2}\,dt$$

for arc-length is standard in textbooks but is often given without rigorous justification or under needlessly restrictive conditions. In this note it is shown that, within the context of the Lebesgue integration and differentiation theory, formulae (Theorems 2 and 3) can be given for the arc-length which are valid for *any* rectifiable parametrized curve. In particular, (A) is valid if and only if all the $f_i(t)$ are absolutely continuous with respect to Lebesgue measure on t. The relevant measure and integration theory can be found in Rudin [1, Chapters 1, 2, 6, 8].

DEFINITIONS. Let G be as above. The *arc-length* of G from $t=0$ to $t=x\leqq c$ is defined by

$$\text{(B)} \qquad s(x)=\sup_P \sum_{j=1}^{m} \sqrt{\sum_{i=1}^{n} (f_i(t_j)-f_i(t_{j-1}))^2}$$

over all partitions $P: 0=t_0<t_1<\cdots<t_m=x$ of $[0,x]$, provided that this exists finitely. G is *rectifiable* if $s(c)$ exists.

Recall that the *total variation* of a real function $f(t)$ defined on $[0,x]$ is by definition $v(x)=\sup_P \sum_{j=1}^m |f(t_j)-f(t_{j-1})|$ over partitions P of $[0,x]$. If $v_1(x),\dots,v_n(x)$ are the total variations of $f_1(t),\dots,f_n(t)$ then the inequality

$$n^{-\frac{1}{2}}\sum_1^n |x_i| \leqq \sqrt{\sum_1^n x_i^2} \leqq \sum_1^n |x_i| \quad \text{implies} \quad n^{-\frac{1}{2}}(v_1(x)+\cdots+v_n(x)) \leqq s(x) \leqq (v_1(x)+\cdots+v_n(x)).$$

Hence G is rectifiable if and only if all the co-ordinate functions $f_i(t)$ are of *bounded variation* in $[0,c]$ and from now on this will be assumed.

THEOREM 1. *Let G^* be the curve $(v_1(t),\dots,v_n(t))$ obtained by replacing the co-ordinate functions of G by their total variations. If $s^*(x)$ is the arc-length function for G^* then $s^*(x)=s(x)$.*

Proof. By a standard result on functions of bounded variation, provable, e.g. from [1, Theorem 8.14, p. 163], the continuity of the f_i implies that of the v_i so that G^* is a continuous curve.

Since $|f_i(t_j)-f_i(t_{j-1})|\leqq|v_i(t_j)-v_i(t_{j-1})|$ it follows from (B) that $s(x)\leqq s^*(x)$. Now let $0\leqq a<b\leqq c$ and $P: a=a_0<a_1<\cdots<a_m=b$ be any partition of $[a,b]$. Using Minkowski's inequality,

$$\begin{aligned} s(b)-s(a) &= \sup_P \sum_{j=1}^{m} \sqrt{\sum_{i=1}^{n} (f_i(a_j)-f_i(a_{j-1}))^2} \\ &\geqq \sup_P \sqrt{\sum_{i=1}^{n}\left(\sum_{j=1}^{m} |f_i(a_j)-f_i(a_{j-1})|\right)^2} = \sqrt{\sum_{i=1}^{n} (v_i(b)-v_i(a))^2}. \end{aligned}$$

Applying this inequality in intervals of partitions of $[0,x]$ and summing, it follows that $s(x) \geqq s^*(x)$ and hence $s^*(x) = s(x)$. Q.E.D.

The value of this theorem is that the $v_i(t)$ are continuous monotonic increasing functions, which are easier to deal with.

Associated with any continuous monotonic increasing real f on $[0,c]$ with $f(0)=0$ there is a positive Borel measure μ on $[0,c]$ defined by $\mu([0,x)) = f(x)$ — see [1, p. 163]. Let $\mu_1, \ldots, \mu_n, \mu_s$ be the measures associated respectively with $v_1, \ldots, v_n, s$ and let μ be any positive Borel measure such that $\mu_i \ll \mu$, $i = 1, 2, \ldots, n$. For example take $\mu = \mu_1 + \cdots + \mu_n$. Then also $\mu_s \ll \mu$ so that by the Radon–Nikodým theorem the derivatives $d\mu_i/d\mu = g_i(x)$ and $d\mu_s/d\mu = g(x)$ are well defined. Our main object is to prove that $g^2 = \Sigma_{i=1}^n g_i^2$, or equivalently that $\mu_s(E) = \int_E \sqrt{\Sigma_1^n (d\mu_i/d\mu)^2}\, d\mu$ for μ-measurable sets E.

Lemma 1. *Let (X, μ) be a measure space and g_i, $i = 1, 2, \ldots, n$ be non-negative measurable functions on X. Then* $\sqrt{\Sigma_1^n (\int_E g_i d\mu)^2} \leqq \int_E \sqrt{\Sigma_1^n g_i^2}\, d\mu$.

Proof. Suppose first that g_i are simple functions mutually constant on sets $M_1, \ldots, M_k$ of a partition of X (for definitions and theory see [1, Chapter 1]). Then, using Minkowski's inequality again,

$$\sqrt{\sum_1^n \left(\int_E g_i d\mu\right)^2} \leqq \sum_{j=1}^k \sqrt{\sum_{i=1}^n \left(\int_{E\cap M_j} g_i d\mu\right)^2} = \sum_{j=1}^k \int_{E\cap M_j} \sqrt{\sum_{i=1}^n g_i^2}\, d\mu = \int_E \sqrt{\sum_1^n g_i^2}\, d\mu.$$

This proves the lemma for simple functions and the result follows for measurable g_i by taking suitable limits $s_r^i \uparrow g_i$ of simple functions and using the Lebesgue monotone convergence theorem.

Lemma 2. *Let (X, μ) be a measure space and g, g_i, $i = 1, \ldots, n$ be non-negative measurable functions on X. Then if $(\int_E g d\mu)^2 \geqq \Sigma_1^n (\int_E g_i d\mu)^2$ for all measurable $E \subseteq X$, it follows $g^2 \geqq \Sigma_1^n g_i^2$ a.e. $[\mu]$.*

Proof. If $g_1, \ldots, g_n$ are simple and mutually constant on sets $M_1, \ldots, M_k$ of a partition of X then taking $E \subseteq M_j$ gives, $\int_E g d\mu \geqq \mu(M_j \cap E) \cdot \sqrt{\Sigma_1^n g_i^2}$, where the expression under the root is to be evaluated at any point of M_j, being constant there. This forces $g \geqq \sqrt{\Sigma_1^n g_i^2}$ a.e. in M_j and so a.e. in X. For general g_i apply what has been proved to g and a sequence of simple functions $s_r^i \uparrow g_i$.

Theorem 2. *Let $\mu_1, \ldots, \mu_n, \mu_s$ be defined for the rectifiable curve G^* as above. Then $\mu_s(E) = \int_E \sqrt{\Sigma_1^n (d\mu_i/d\mu)^2}\, d\mu$ for μ-measurable E and in particular $s(x) = \int_{[0,x]} \sqrt{\Sigma_1^n (d\mu_i/d\mu)^2}\, d\mu$.*

Proof. It is sufficient to check the assertion for intervals E, the extension to measurable E following by the regularity of all measures concerned [1, pp. 41, 48]. With $E = [a, b]$ and partitions $P: a = a_0 < a_1 < \cdots < a_m = b$,

$$\mu_s(E) = \mu_s([a,b]) = \sup_P \sum_{j=1}^m \sqrt{\sum_{i=1}^n \mu_i([a_{j-1}, a_j])^2} = \sup_P \sum_{j=1}^m \sqrt{\sum_{i=1}^n \left(\int_{[a_{j-1}, a_j]} g_i d\mu\right)^2}$$

$$\leqq \sup_P \sum_{j=1}^m \int_{[a_{j-1}, a_j]} \sqrt{\sum_{i=1}^n g_i^2}\, d\mu = \int_E \sqrt{\sum_1^n g_i^2}\, d\mu,$$

using Lemma 1. Again, $\mu_s(E)^2 \geqq \Sigma_1^n \mu_i(E)^2$ for intervals and so for general E by regularity. Using Lemma 2, $g^2 \geqq \Sigma_1^n g_i^2$ a.e. $[\mu]$ and the reverse inequality $\mu_s(E) \geqq \int_E \sqrt{\Sigma_1^n g_i^2}\, d\mu$ follows on integrating. This completes the proof. Q.E.D.

In general, it is not possible in Theorem 2 to take μ as Lebesgue measure m; however, we can split the integral there into the sum of a Lebesgue integral and an integral with respect to a mutually singular measure $\lambda \perp m$ by using the Radon–Nikodým theorem [1, pp. 121, 122].

Decompose $v_i(t)$ into an absolutely continuous part and a singular part [1, p. 166]: $v_i(t) = p_i(t) + q_i(t)$ with $p_i(t) = \int_0^t v_i'(y)dy$ and $q_i'(t) = 0$ a.e. $[m]$ and p_i, q_i continuous monotonic increasing. Let σ_i, λ_i be the positive Borel measures associated with p_i, q_i respectively so that $\mu_i = \sigma_i + \lambda_i$, $\sigma_i \ll m$, $\lambda_i \perp m$. Put $\lambda = \sum_1^n \lambda_i$ and take $\mu = m + \lambda$, so that $\lambda_i \ll \lambda$, $\mu_i \ll \mu$ and $\lambda \perp m$.

LEMMA 3. *Let (X, μ) be a measure space and suppose $\mu = \nu + \lambda$ where $\nu \perp \lambda$. Suppose also $\sigma \ll \nu$, $\lambda' \ll \lambda$ and $\mu' = \sigma + \lambda'$ (all measures positive). Then if $f: X \to R$ is integrable $[\mu]$ over X and $E \subseteq X$ is μ-measurable, $\int_E f d\mu = \int_E f d\nu + \int_E f d\lambda$, $d\mu'/d\mu = d\sigma/d\nu$ a.e. $[\nu]$, $d\mu'/d\mu = d\lambda'/d\lambda$ a.e. $[\lambda]$.*

Proof. This is an elementary consequence of the definitions and the uniqueness part of the Radon–Nikodým theorem, and the details are omitted.

THEOREM 3. *Let G_1 be the curve $(p_1(t), \ldots, p_n(t))$ and G_2 the curve $(q_1(t), \ldots, q_n(t))$ defined from the decomposition $v_i = p_i + q_i$, with respective arc-length functions $s_1(x)$ and $s_2(x)$. Then $s(x) = s_1(x) + s_2(x)$ and*

$$s_1(x) = \int_0^x \sqrt{\sum_1^n p_i'(t)^2}\, dt = \int_0^x \sqrt{\sum_1^n f_i'(t)^2}\, dt$$

$$s_2(x) = \int_{[0,x]} \sqrt{\sum_1^n (d\lambda_i/d\lambda)^2}\, d\lambda.$$

Proof. Apply Lemma 3 to the integral in Theorem 2 with $\nu = m$ as Lebesgue measure to obtain a sum of two integrals, and then use the second part of the lemma to replace $d\mu_i/d\mu$ by $p_i'(t) = d\sigma_i/dm$ in the Lebesgue integral and by $d\lambda_i/d\lambda$ in the λ-integral.

That $|f_i'(t)| = v_i'(t) = p_i'(t)$ a.e. $[m]$ is a standard result about functions of bounded variation — provable, e.g., by considering the absolutely continuous parts of the complex measure α associated with $f_i(t)$ and of $|\alpha|$ and using (1, 6.13, 8.14, 8.18). As an immediate corollary, formula (A) is valid just when the co-ordinate functions $f_i(t)$ are all absolutely continuous. Q.E.D.

Example: Let G be a plane curve $(t, f(t))$ with graph $y = f(x)$. Then $p_1(t) = t$, $\lambda_1 = 0$ so that

$$s(x) = \int_0^x \sqrt{1 + f'(t)^2}\, dt + \lambda_2([0, x]) = \int_0^x \sqrt{1 + f'(t)^2}\, dt + v(x) - \int_0^x |f'(t)|\, dt.$$

In particular, $s(x) \leqq x + v(x)$ with equality if and only if $f'(t) = 0$ a.e. in $[0, x]$ — c.f. Problems 6007 and 6074, this MONTHLY, 82 (1975) 84, and 83 (1976) 140.

Acknowledgement. The author wishes to thank the referee for a number of suggestions which improved the exposition of this note.

Reference

1. W. Rudin, Real and Complex Analysis, McGraw-Hill, New York, 1970.

164 BALMORAL ROAD, WATFORD, HERTFORDSHIRE, WD2 4EY, ENGLAND.

A SIMPLIFICATION IN CERTAIN CONTOUR INTEGRALS

HAROLD P. BOAS AND EDUARDO FRIEDMAN

It appears to have escaped notice in the literature that the evaluation by residues of definite integrals containing an exponential factor can be simplified by use of a triangular contour. Consider for example $\int_{-\infty}^{\infty} R(x)e^{ix}dx$ where $R(x)$ is a rational function vanishing at infinity. We integrate $R(z)e^{iz}$ over the triangle with vertices at $-\rho_1$, ρ_2, and $(\rho_1 + \rho_2)i$, where $\rho_1 > 0$ and $\rho_2 > 0$ are taken sufficiently large that the contour encloses all the poles in the upper half-plane. By the hypothesis

$|R(z)|$ is bounded by a constant times $1/|z|$, which is bounded by $\sqrt{2}/\rho_2$ on the diagonal line in the first quadrant. Since $|dz| = (1+\rho_2^2/(\rho_1+\rho_2)^2)^{\frac{1}{2}}dy \leqq \sqrt{2}dy$, it follows that the contribution from this line is less than a constant times $(1/\rho_2)\int_0^{\rho_1+\rho_2} e^{-y}dy \leqq 1/\rho_2$. Similarly the contribution from the line in the second quadrant is less than a constant times $1/\rho_1$. By letting $\rho_1, \rho_2 \to \infty$ it follows that the integral over the real axis equals $2\pi i$ times the sum of the residues in the upper half-plane. This estimate is simpler than the computation with either a semi-circular or rectangular contour because the slope of the triangular contour is bounded away from zero in the upper half-plane.

As a second example we establish the formulas

$$\int_0^\infty x^{s-1}\cos x dx = \Gamma(s)\cos\tfrac{1}{2}\pi s, \quad \int_0^\infty x^{s-1}\sin x dx = \Gamma(s)\sin\tfrac{1}{2}\pi s, \qquad 0<s<1,$$

which are the real and imaginary parts of $\int_0^\infty x^{s-1}e^{ix}dx = \Gamma(s)e^{\frac{1}{2}i\pi s}$. Usually [1, p. 107] one integrates $z^{s-1}e^{iz}$ over a contour consisting of the real axis from δ to ρ, the circle $|z| = \rho$ from $\theta = 0$ to $\theta = \frac{1}{2}\pi$, the imaginary axis from $i\rho$ to $i\delta$, and the circle $|z| = \delta$ from $\theta = \frac{1}{2}\pi$ to $\theta = 0$. As $\delta \to 0$ the contribution from the arc of $|z| = \delta$ vanishes like δ^s. No poles are enclosed by the contour, so

$$\int_0^\infty x^{s-1}e^{ix}dx = \int_0^{i\infty} z^{s-1}e^{iz}dz = \int_0^\infty (iy)^{s-1}e^{-y}idy = e^{\frac{1}{2}i\pi s}\int_0^\infty y^{s-1}e^{-y}dy = e^{\frac{1}{2}i\pi s}\Gamma(s).$$

This result holds assuming that the contribution from the large quarter circle vanishes as $\rho \to \infty$.

As before, we save work in making this estimate by connecting the points $z = \rho$ and $z = i\rho$ by a straight line, rather than by the arc of a circle. On this line, $|dz| = \sqrt{2}dy$ and $1/|z| \leqq \sqrt{2}/\rho$. It follows immediately that the integral over the line is bounded by a constant times $\rho^{s-1}\int_0^\rho e^{-y}dy \leqq \rho^{s-1}$ which tends to zero as $\rho \to \infty$ since $s < 1$.

Reference

1. E. C. Titchmarsh, The Theory of Functions, 2nd ed. Oxford University Press, London, 1958.

CORRESPONDENCE ADDRESS: HAROLD P. BOAS, GRADUATE MATHEMATICS OFFICE, MIT, CAMBRIDGE, MA 02139.

A THREE-DIMENSIONAL SYSTEM WITH KNOTTED TRAJECTORIES

RICHARD PARRIS

A detailed look at autonomous systems of first-order equations in the plane is an enjoyable part of any introductory course in differential equations, especially for those teachers and students who like to sketch solutions. The following example gives an effective indication of the complexity possible in systems of dimension greater than two. And, although it is nonlinear, the example is simple enough to appear uncontrived.

The system is

$$\begin{aligned} x' &= -m\cdot y + n\cdot x\cdot z \\ y' &= m\cdot x + n\cdot y\cdot z \\ z' &= (n/2)\cdot[1+z^2-x^2-y^2]. \end{aligned}$$

The solution that passes through (a, b, c) when $t = 0$ is

$$x = \frac{2\cdot a\cdot\cos(m\cdot t) - 2\cdot b\cdot\sin(m\cdot t)}{\Delta - 2\cdot c\cdot\sin(n\cdot t) + (2-\Delta)\cdot\cos(n\cdot t)},$$

$$y = \frac{2 \cdot a \cdot \sin(m \cdot t) + 2 \cdot b \cdot \cos(m \cdot t)}{\Delta - 2 \cdot c \cdot \sin(n \cdot t) + (2 - \Delta) \cdot \cos(n \cdot t)},$$

$$z = \frac{2 \cdot c \cdot \cos(n \cdot t) + (2 - \Delta) \cdot \sin(n \cdot t)}{\Delta - 2 \cdot c \cdot \sin(n \cdot t) + (2 - \Delta) \cdot \cos(n \cdot t)},$$

where Δ denotes $1 + a^2 + b^2 + c^2$. Unless $a^2 + b^2 = 0$, this trajectory lies on the torus obtained by revolving the circle $(x - K)^2 + z^2 = K^2 - 1$ about the z-axis, where K stands for $\Delta/(2\sqrt{a^2 + b^2})$. The z-axis is an exceptional trajectory; so too is the unit circle in the xy-plane, which corresponds to the degenerate torus $K = 1$.

Establishing the preceding assertions requires computation, of course, but the work is not tedious if it is economically organized. One way of proceeding is to introduce extra notation:

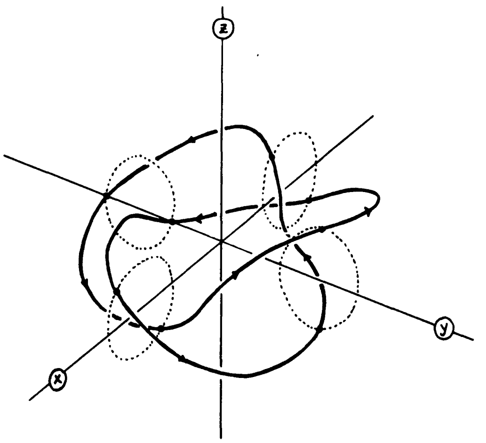
$$x = X/D, \quad y = Y/D, \quad z = Z/D,$$

where D denotes $\Delta - 2 \cdot c \cdot \sin(n \cdot t) + (2 - \Delta) \cdot \cos(n \cdot t)$. Then one may establish the relationships $X' = -m \cdot Y$,

$$Y' = m \cdot X, \quad Z' = n \cdot [D - \Delta], \quad D' = -n \cdot Z, \quad X^2 + Y^2 = 4a^2 + 4b^2,$$

and $D^2 + X^2 + Y^2 + Z^2 = 2 \cdot \Delta \cdot D$. The desired conclusions follow easily from these formulas.

If m and n are coprime integers, the unexceptional trajectories are all torus knots of type (m, n), winding m times about the z-axis while winding n times about the unit circle. The trajectories are pairwise linked. The diagram shows a representative trajectory in the case $m = 2$, $n = 3$. It is a trefoil knot.



If m and n are incommensurable, the unexceptional trajectories are not closed curves, and in fact are dense in their tori.

It is also interesting to see the effect of the nonlinear terms by studying the related linear system, whose solutions are helices.

Department of Mathematics, Princeton University, Princeton, NJ 08540.

A NOTE ON THE CENTRAL LIMIT THEOREM

CHI SONG WONG

Khan [2] has pointed out that giving a short proof of Stirling's formula is an attractive application of the limit theorems which are a standard part of courses in probability at the graduate level. He gave such a proof by applying the central limit theorem and a moment convergence theorem to exponentially distributed random variables. Here we give a much shorter proof of Stirling's formula by applying these tools to a function of random variables having a Poisson distribution.

Historically, Stirling's formula was used to derive the central limit theorem for independent Bernoulli's (coin tossing) random variables X_n's of parameter p, where $S_n \equiv X_1 + X_2 + \cdots + X_n$ is then a Binomial random variable of parameter (n, p). When p is small and n is large, Poisson distribution of parameter $\lambda = np$ is used to approximate Binomial distribution of parameter (n, p) [3]. This motivates our use of Poisson random variables to derive Stirling's formula: Let $\{X_n\}$ be a sequence of independent Poisson random variables of parameter 1 on some probability space. Then $S_n = X_1 + X_2 + \cdots + X_n$ is Poisson of parameter n. So the standardized random variable of S_n is $(S_n - n)/\sqrt{n}$. For any real-valued function f on a set Ω, let f^- denote the function on Ω defined by $f^-(w) = -f(w)$ if $f(w) < 0$ and $f^-(w) = 0$ if $f(w) \geqq 0$. Then

$$E\left(\left(\frac{S_n - n}{\sqrt{n}}\right)^-\right) = e^{-n} \sum_{j=0}^{n} \left(\frac{n-j}{\sqrt{n}}\right) \frac{n^j}{j!} = \frac{\sqrt{n}(n/e)^n}{n!}$$

(all terms but one are cancelled). Let Z be a standardized normal random variable. By the central limit theorem [1, Theorem 6.44, p. 169], $\{(S_n - n)/\sqrt{n}\}$ converges in distribution to Z. Let g be the identity function on the real line. Since g^- is continuous,

$$\left\{\left(\frac{S_n - n}{\sqrt{n}}\right)^-\right\}$$

converges in distribution to Z^- [4, Theorem 3, p. 90]. Thus by the moment convergence theorem [1, Theorem 4.52, p. 95],

$$\lim_{n \to \infty} E\left(\left(\frac{S_n - n}{\sqrt{n}}\right)^-\right) = E(Z^-).$$

Since $E(Z^-) = 1/\sqrt{2\pi}$,

$$\lim_{n \to \infty} \frac{\sqrt{n}(n/e)^n}{n!} = \frac{1}{\sqrt{2\pi}}.$$

Hence

$$\lim_{n \to \infty} \frac{\sqrt{2\pi n}(n/e)^n}{n!} = 1.$$

This research was supported by National Research Council of Canada Grant No. A8518.

References

1. K. L. Chung, A course in probability theory, 2nd ed., Academic Press, New York, 1974.
2. R. A. Khan, A probabilistic proof of Stirling's formula, this MONTHLY, 81 (1974) 366–369.
3. L. E. Maistrov, Probability theory, a historical sketch, Academic Press, New York, 1974.
4. H. G. Tucker, A graduate course in probability, 2nd printing, Academic Press, New York, 1971.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WINDSOR, WINDSOR, ONTARIO, CANADA N9B 3P4.

MATHEMATICAL EDUCATION

EDITED BY PAUL T. MIELKE AND SHIRLEY HILL

Material for this Department should be sent to Paul T. Mielke, Department of Mathematics, Wabash College, Crawfordsville, IN 47933, or to Shirley Hill, Department of Mathematics, University of Missouri, Kansas City, MO 64110.

THE APPRENTICE SYSTEM FOR MATH METHODS

J. A. MURTHA

"The best way to learn is to do; the worst way to teach is to talk." Paul Halmos, this MONTHLY May 1975.

Introduction. Marietta College is one of several hundred predominantly undergraduate institutions that offers certificates for teachers of high school mathematics. Each of these colleges and universities must offer a course in methodology usually called something like "Methods of Teaching: Mathematics." Having taught such a course for a few years and having discussed the subject with several colleagues from other schools, I can only assume that many of us have been less than satisfied with the fruits of our labors. This paper outlines a non-traditional approach that utilizes the students as apprentices to college instructors.

The traditional methods course at Marietta. For several years now, the math methods course has been offered once each year by the department of mathematics. It carries two semester hours credit and complements a two credit hour "general methods" course offered by the education department. Our students are usually a mixture of juniors and seniors, some of whom have already done student teaching at a local school. Most but not all of them are math majors. Nearly every year we have a psychology major (with a math minor), a physics major or some other science majors. In Ohio, as in most states, it is not necessary to have majored in a subject in order to teach it: one simply must obtain credit in a minimum number of courses in that discipline. The methods class is small, ranging from five to twelve students during the past six years.

In the past, the course had a textbook with assigned readings, but the text never played a central role in the course. The course featured presentations to the class by individual students, the design of visual aids, group discussions on some potentially controversial topics, practice at making up quizzes and exams, and a term paper with very broad guidelines. For the last four years, each student gave one lecture to an elementary college level class, usually after presenting that lecture to and getting a thorough critique from the methods class. In addition, the student would present another lecture to the methods class on a topic not typically encountered in their other undergraduate math courses.

Weaknesses in the traditional course. Over a period of time, the individual presentations became the most significant aspect of the course in the eyes of the students. It was agreed that, except for the dry run of the college class lecture, all presentations had to fit the actual audience rather than, say, giving a 10th grade lecture to a group of college students. We wanted a sense of realism as well as an atmosphere conducive to spontaneous dialog between lecturer and audience. Of the twenty-odd talks to our college courses that I have witnessed, none was embarrassing. The methods students viewed their assignment seriously and spent abundant time in preparation.

As it turned out, each student in the methods class found the experience of lecturing to students who were practically his or her peers an exhilarating experience. It overshadowed everything else in the course. They could not get excited preparing quizzes or worrying about grading policies or planning material for the next few weeks or producing visual aids. Their experience with that class, unfortunately, was limited to one lecture. They had no vested interest in the overall success of the course. They were but one of several students making cameo appearances.

Thus, despite the fact that each student became deeply interested in teaching mathematics for a brief period preceding his or her two performances, for the most part the class seemed unmotivated and unwilling to contribute. They would gear up once again to write a paper near the end of the semester, but both they and the instructor usually came away from the course feeling it ought to have been more productive.

The missing ingredient — long term commitment. Looking back now, it seems obvious why the experience always seemed frustrating. In our efforts to approximate reality we were simply too myopic. As any teacher knows, it is not the single day, good or bad, that determines the level of success of a given course. On the contrary, teaching effectiveness ought to be measured by the extent to which all the students are changed, not only in terms of knowledge but in terms of attitude toward the subject and toward learning in general. An important part of teaching is being able to adjust one's style and one's plans to the needs and abilities of a particular group of individuals. When a student invests a year or a half-year of time in a course, then the teacher deserves to be praised or blamed for having some definite effect on that student's intellectual growth. It is precisely this factor of time, of long-term involvement, that makes a teacher accountable, thereby providing a large incentive to do the job well. Our students in the methods class rightly perceived that they were having little actual effect on the class to whom they were giving one lecture. Once the thrill of the presentation wore away, they felt cheated and lost interest in the freshman course as well as the methods course itself.

The non-traditional experiment. Largely by accident in the spring of 1974 we discovered a way to present a prospective high school teacher with the full range of joy and frustrations associated with teaching a complete course. One of our seniors wanted to study geometry in a semester when the appropriate (advanced) course was not offered. We arranged that she become an assistant to one of our instructors who was teaching a geometry course to prospective elementary school teachers. The course seemed well-suited to having an assistant because it was utilizing a small group discovery method instead of lectures. The assistant's primary responsibility was to facilitate the group process, aiding or observing one group while the instructor worked with another group. Naturally, to be effective in this role she had to understand thoroughly the material in the text used by the class.

A second responsibility of the assistant was to investigate alternate approaches to the subject by perusing other texts and then periodically discussing them with the instructor. These discussions led ultimately into a broad range of pedagogical issues. By the end of the semester, the assistant had participated in a number of activities, including preparation and grading of quizzes. A final report written by the assistant featured a critique of the course and some generalizations about student attitudes toward the course, as well as a factual account of her activities and an evaluation of the same. It was clear to all concerned that the experiment had been successful.

Development of the apprentice system for the math methods course. During the fall of 1974 we laid the groundwork for a methods course fashioned after the experiment of the preceding spring. We identified the students who would take the methods course in the spring of 1975. Five students were signed up: two senior math majors who were currently doing student teaching and three juniors, one of whom was a psychology major with a strong minor in mathematics. Five faculty members volunteered to accept one student each as an "apprentice," the name we felt most suitable for the roles they would play. We discussed the variety of activities that these apprentices might engage in. Meanwhile we notified the prospective apprentices about our plans and asked them to give some thought to the course and the instructor with whom they would like to be associated.

Both faculty and students were forewarned of the danger of converting the apprentice into a glorified paper-grader. From the outset, our intent was to devise a broad learning experience for the apprentices. They were not to become lackeys; they were to be placed in a position allowing them to observe how an instructor plans and teaches a course.

By the time the spring semester began we had paired up the apprentices with faculty "mentors." After one organizational meeting the apprentices began working with their assigned courses. The

ground rules were fairly simple: I would meet with all the apprentices approximately once each week for the first two or three weeks and once every two weeks thereafter; prior to our meetings each apprentice would fill out and submit for the rest of us to read a one page questionnaire indicating how he had been spending his time, what his plans were, how useful he felt and what topics he wanted to discuss at our meeting. Students retained copies of these progress reports and used them as resource material when they wrote their final report. In the final reports, students described their activities in detail, assessed their learning experience as well as their perception of how they were viewed by their mentor and by the students in the class in which they served. They concluded with recommendations for improvements in the overall program.

At our organizational meeting I enumerated several activities that seemed appropriate for the apprentices: introducing and directing classroom activities (especially in our course in liberal arts mathematics); preparing enrichment materials (for example, a supplementary lecture, a reading list or a collection of small group or individual projects); preparing study sheets for examinations; preparing audio-visual material; conducting study sessions for exams or even on a routine basis; facilitating small group activities (especially in the geometry course for prospective elementary school teachers); examining alternative textual material and discussing this as well as alternative methodologies with the mentor; preparing and grading quizzes and exams; leading the classroom discussion of the preceding day's homework assignment (which turned out to be a good warm-up for actual lecturing); delivering regular class lectures; serving as an (objective) classroom observer and giving the mentor feedback on his classroom effectiveness.

Needless to say, no apprentice did all these things. Instead, each apprentice-mentor team chose a suitable subset of these activities. At least twice in each course, the apprentice had the full responsibility for planning and presenting the lecture (or heading the discussion in the case of the geometry class).

Our bi-weekly meetings were initially devoted to sharing of experiences and planning activities. After the roles became better identified, we discussed a variety of topics including motivating students; the possible roles of homework, quizzes, exams, classroom discussion, as they relate to determining a course grade; "the things I most (dis-)liked about some of my mathematics instructors"; classroom activities, projects and games; Paul Halmos', "Mathematics as a Creative Art" from *American Scientist*; and professional organizations and keeping mathematically alive after undergraduate school.

Conclusions. Without exception all the apprentices and mentors felt that the experience was valuable for the apprentices, the mentors and the test classes. The apprentices strongly urged that this program permanently replace our traditional methods course. They generated some promising suggestions for modifications. In particular, we all felt that the perspective of the apprentice as a classroom observer allowed him to regard the instructor in an interesting way. It became a study in pedagogical style and the mentor served as one model. Even though the apprentice may have taken courses from that mentor, all of a sudden the relationship was changed. The apprentices recommended that more opportunity be made available to their successors for observing other instructors in order to compare styles. We hope to arrange a series of discussions of style, perhaps involving role playing, that would feature the college instructors as models for the prospective teachers.

From the mentor's the experiment forced him to articulate some of his teaching philosophy. For example, when he and his apprentice would discuss a quiz before it was given and then discuss the grading of it afterwards, it became necessary to explain what was being tested, why and how it was being done, why partial credit may or may not be appropriate, and how the grade was being used.

The students in the classes were overwhelmingly in favor of the apprentices. They generally received more attention than in other courses and did not feel that the course suffered in any way. Some students went so far as to credit the apprentice for their success in the course.

Beyond this alleged satisfaction of the three parties lies a serious question, called to the author's

attention by Professor Zalman Usiskin who was kind enough to read an earlier version of this paper: teaching college students is quite a different task from teaching high school students; to what extent does the apprentice experience prepare one to teach in high school?

Given the difference in environment, responsibilities and expectations there seem to be some basic principles common to teaching at all levels. To be effective, any teacher must, among other things, have the ability to (1) plan and organize material, (2) convey a sense of enthusiasm toward the subject, (3) evaluate the learners' progress, and (4) cope with a range of personalities. It seems to me that any kind of experience in these fundamentals is not only valuable but transferable from one level of teaching to another.

Perhaps an even better opportunity would exist at institutions having an affiliated "university high school" where a faculty member (mentor) and his or her apprentice could actually team up to teach a high school class. Short of that, however, I think much can be gained with the system as described here.

DEAN OF THE COLLEGE, MARIETTA COLLEGE, MARIETTA, OH 45750.

TUTOR RESOURCES IN MATHEMATICS

LEONARD SHAPIRO

Consider the math student who has been taught some material in a course, by one or more of the usual teaching instruments (text, lecture, T.V., computers, etc.). This student then tries to solve the homework problems. Often, he or she cannot. Even after reviewing the text, notes, etc., and listening to what is said in class, help is still needed. Most schools supply some kind of resources for the student in this situation, including complete printed solutions to problems, an instructor's office hours, or student tutors for that course or several math courses. For lack of a better term, let us call these "tutor resources."

This paper describes a new kind of tutor resource called ISCI (Individualized Supplementary Calculus Instruction). Studies we have conducted show that ISCI is at least as helpful as traditional tutor resources, and that for weaker students it is more helpful.

Tutor resources are not the most important component of a course, but to some students they are invaluable. On the other hand, many students will never use them. In fact, they are often too little used by the students who most need help. In this note we will argue that different resources, including ISCI, appeal to different students and serve different functions. Thus, making available to students a wide variety of resources, including ISCI, should be preferable to reliance on a few.

Although this paper is mainly intended to encourage implementation of a resource like ISCI, I hope that it can also shed some light on the question of why the traditional tutor resources do or do not help students. ISCI is a very different kind of resource, and in comparing its qualities to those of the traditional resources we may see why one or the other is more helpful to particular students. This may enable us to improve existing resources and also to guide each student to the particular one that will suit that student best.

The ISCI Program. ISCI is a computer-based resource that has been in use since fall, 1973, at several campuses in Minnesota, Wisconsin and Illinois, on Control Data, IBM and Univac computer systems. The ISCI program has been designed to be used with any text or computer system, and once installed it needs no maintenance. Development of ISCI has been supported by a grant from the Fund for the Improvement of Postsecondary Education, an HEW agency.

In brief, ISCI works in the following way: it assumes that the student has tried to solve certain homework problems, and failed. If the student chooses to use ISCI, it will ask him (through a teletype

or other computer terminal) about his attempted solution. After finding the error in his attempted solution, ISCI points out that there is an error, then gives as many hints and/or explanations as the student requests, until he is able to continue the solution independently.

The dialogue below is a sample of what ISCI looks like to a student. Underlined material was typed by the student. The problem was to find the local maxima and minima of $f(x) = x^4 - 2x^2$.

WHAT PROBLEM ARE YOU HAVING TROUBLE WITH?
? MM4
WHERE WAS YOUR TROUBLE?
A. I COULDN'T FIND $F'(X)$.
B. I FOUND THE DERIVATIVE, BUT I COULDN'T FIND THE CRITICAL POINTS.
C. I FOUND THE CRITICAL POINTS, BUT I COULDN'T DETERMINE WHICH WERE LOCAL MAXIMA OR MINIMA.
D. I COULDN'T DETERMINE WHICH OF THE LOCAL MAXIMA AND MINIMA WERE ABSOLUTE MAXIMA AND MINIMA.
E. MY ANSWERS DON'T AGREE WITH THE BOOK'S ANSWERS.
F. I DIDN'T GET AN ANSWER.
? E
LET'S SEE WHICH ONE OF YOU HAS THE RIGHT ANSWERS.

THE CRITICAL POINTS I FOUND WERE
A. 0, 1
B. -1, 0, 1
C. -1. 1
D. 0, SQRT 2, $-$SQRT 2
E. MY ANSWER DOES NOT APPEAR
? D
THERE'S AN ERROR IN YOUR SOLUTION HERE. ENTER YES FOR A HINT ABOUT YOUR ERROR, NO TO TRY THE QUESTION AGAIN.
? YES
HINT: THE CRITICAL POINTS ARE THE POINTS X SUCH THAT $F'(X) = 0$, NOT $F(X) = 0$.
ENTER NO TO TRY AGAIN, YES TO GET AN EXPLANATION OF HOW TO ANSWER THE QUESTION.
? YES
LET'S GO BACK A STEP

WHAT IS $F'(X)$?
A. $X^3 - 2X$
B. $4X^3 - 2X$
C. $4X^3 - 4X$
D. MY ANSWER DOES NOT APPEAR
? A
THERE'S AN ERROR IN YOUR SOLUTION HERE. ENTER YES FOR A HINT, NO TO TRY AGAIN.
? YES
REMEMBER THAT THE DERIVATIVE OF X^N IS NX^{N-1}. DID YOU FORGET THE N?
ENTER NO TO TRY AGAIN, YES TO GET AN EXPLANATION OF HOW TO ANSWER THE QUESTION.
? STOP

Although only single-letter responses (A, B, etc.) are used in this simple example, words or expressions can be used for possible multiple-choice responses. Detailed explanations may be given, one step at a time, if appropriate. This student's mistake was simple and typical.

ISCI is by no means a very sophisticated error-diagnosis mechanism. It is, however, a simple and efficient method for answering the typical "dumb" questions which so often are an embarrassment to students and a time-consuming chore for teachers.

The ISCI package provides help only with homework problems for which "problem files" have been written. These are very simple to write. Most of those presently in use have been written by undergraduates. At present over 100 problem files have been written, all of them for problems in College Algebra and Calculus.

A complete technical description of ISCI is available from the author, along with a copy of the relevant computer program on magnetic tape. There will be no charge for the written materials, but there will be a nominal charge for the tape.

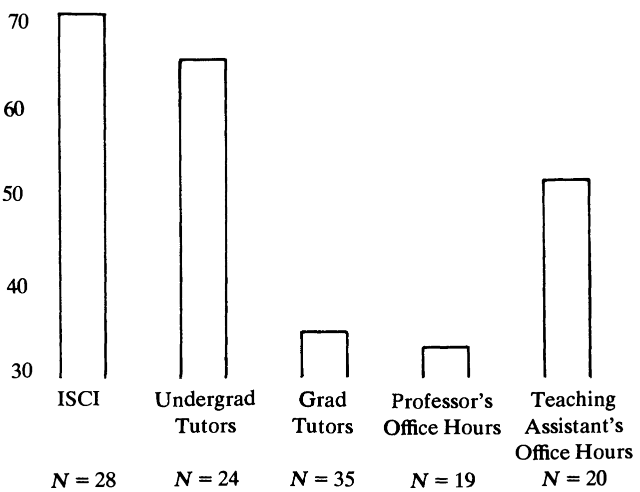
Prospective users are invited to use ISCI verbatim, or to alter or expand either the program or the individual problem files. Several persons with no computer programming experience have used ISCI by asking a local computer programmer to set it up, and then writing their own problem files.

Evaluations of ISCI's effectiveness. A study of ISCI was conducted in 1974, using 400 first-year calculus students from six lecture sections, each under the supervision of a professor. Each of the lecture sections was divided into four or five recitation sections led by teaching assistants; 20 assistants in all were involved. All students had access to ISCI and to four other tutor resources: Professors and T.A.'s during prescribed office hours, and graduate and undergraduate tutors who were available throughout the week. No changes were made in the established curriculum or teaching methods.

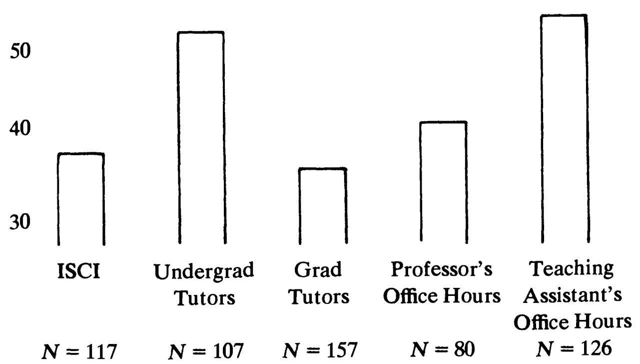
Students were asked to rate the helpfulness of the 11 resources available to them, including the five tutor resources. Since lecture, text and sections were usually ranked 1, 2 and 3, we considered a ranking of 4 or better for a tutor resource to indicate that it was "very helpful." Fig. 1 shows the results. It is clear that each tutor resource is reported as beneficial by some students; in fact, each is ranked very helpful by at least 1/3 of the students who tried it. For students with lowest-quartile grades, resources such as ISCI and undergraduate tutors are much more heavily favored, showing that

FIG. 1. Percentage of users of each resource who rated it as very helpful

(a) Students whose grades in the previous math course were in the lowest quartile ($N = 80$)



(b) Entire population ($N = 407$)

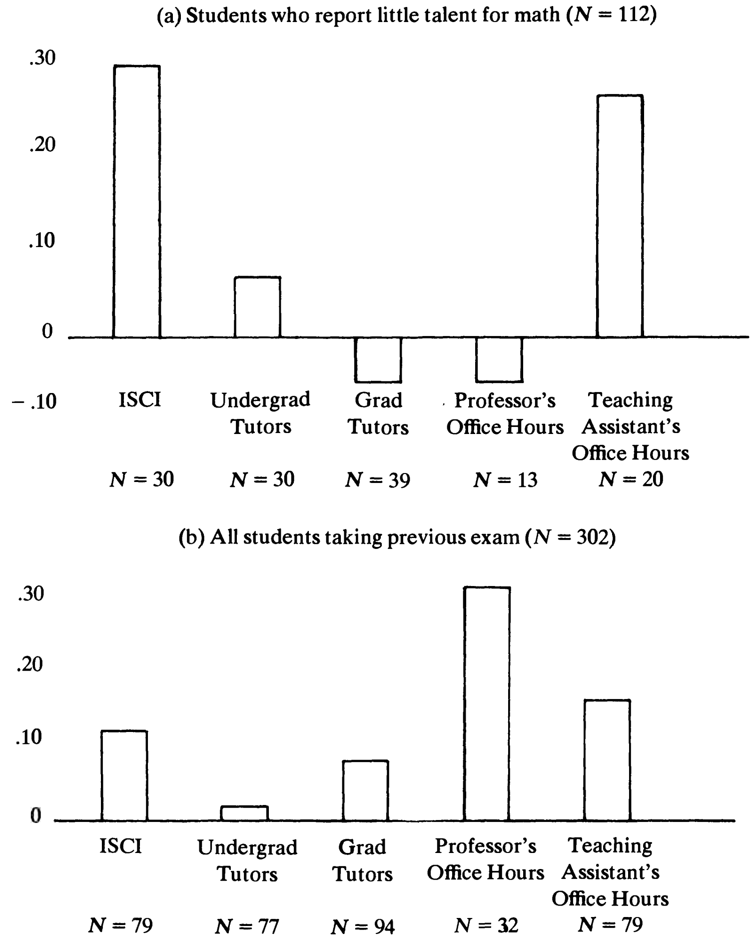


Note: N refers to the number in each group; the groups overlap in students using more than one resource.

these resources are most helpful to these weakest students. In fact, 52% of lowest-quartile students who used ISCI ranked it 3rd or better, while the analogous percentages for other resources were 33% (for undergrad tutors) and below.

Grade changes from the previous quarter to the current one were computed for the 350 students who took the same final exam each quarter. For each resource, grade changes for users of the resource were compared to those for nonusers. Figure 2 displays the differences in grade change for all students and for those with low math aptitude. No significant differences in grade changes of users vs. nonusers were obtained, for ISCI or any other tutor resource, even at a liberal confidence level of $\alpha = .10$.

FIG. 2. Grade Changes for Resource Users Compared to Nonusers

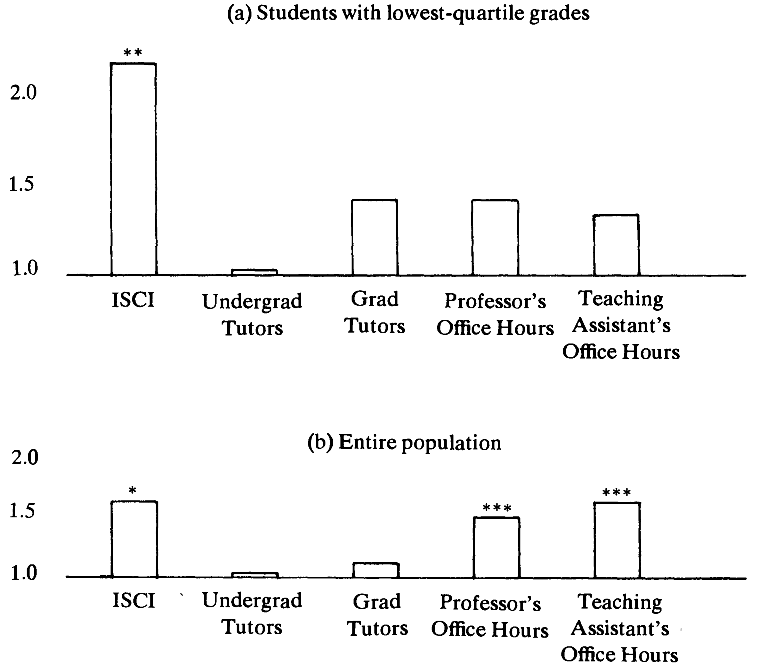


Note: Grade improvement is measured by the difference in (normalized) final exam scores from one quarter to the next. Bars pointing downward occur in situations where nonusers of a resource improved more than did users.

Students were asked how their attitude toward math had changed during the quarter: more favorable, no change, or less favorable, (see Figure 3). Again, for each resource, responses of users of that resource were compared with responses of nonusers. In this case, users of ISCI responded "more favorable" significantly more often ($\alpha = .025$) than did nonusers. This significant difference still holds when we consider only students with lowest-quartile grades in their previous quarter of math.

A more complete analysis of the evaluations we have conducted is available from the author.

FIG. 3. Ratio of percentages of users and of nonusers of each resource, who claimed their attitude toward math had become "more favorable" during the course.



Note: The leftmost bar in the upper graph, with height 2.2, means that if we consider only students with lowest quartile incoming grades, the proportion of students using ISCI who responded "more favorable" attitude was 2.2 times the proportion of nonusers of ISCI who responded "less favorable."
Using original data from which these tables were calculated, the hypothesis: "users and nonusers of a resource were equally likely to claim a more favorable attitude" could be rejected in the four cases marked by asterisks above. A single * refers to an α of .01, ** means $\alpha = .025$, *** corresponds to $\alpha = .05$. For N's see Figure 1.

Qualities of ISCI and other tutor resources. As we have mentioned above, a comparison of the qualities of ISCI and other tutor resources raises interesting and useful questions about which is more helpful to a particular student. There are at least three ways in which ISCI differs significantly from other tutor resources of printed solutions to problems, instructor's office hours, and student tutors.

1. *Personal Remediation*: A computer-assisted program (like a nonhurried and sensitive human tutor) can encourage the student to find his own mistake by following whatever path the student's solution takes and giving judicious hints. Alternatively, the student can just be told what he did wrong, or be shown a correct solution. The latter alternatives are certainly more efficient, but do they teach any more than the ability to imitate? Can allowing the student to find his own error, and the extra time this requires, effectively teach an important aspect of problem-solving?

2. *Accessibility of Resources*: The accesssibility of a resource depends on the individual student as well as on the available facilities. For example, a student may feel so threatened by a human tutor or an instructor (especially if that instructor will be giving the student a grade) that those tutor resources are not acceptable alternatives to that student. Computers do not have this specific disadvantage, but fear of computers in general may keep students away from computer terminals. Facilities that are overcrowded or open very few hours per day can be discouraging, although overcrowding can be used to advantage for some students if they can be encouraged to study together.

But is ease of access an unqualified good? This too depends on the individual student. It has happened to me several times that a student, at the end of a lecture, tells me he has had difficulty solving the problems in a particular topic. I point out where he can get help, including the ISCI program. When I see him again, he will say: "The program (or tutor) was lousy. I tried it but it was so slow (or crowded) that it was faster for me to do it myself." So for this student the necessity to seek out, or wait for, a resource was beneficial: it encouraged him to try harder to do the problem himself. This is not an argument for eliminating all tutor resources — if no aid had been available to the student under any circumstances he may have been too discouraged to continue. Furthermore, the student I am describing is by no means typical, and others cannot be so easily persuaded to do without help.

3. *Range of Material*: Human tutors, ideally, can provide help with any difficulties a student has in a particular course. But this is not the case with printed solutions or ISCI. The latter resources usually cannot easily handle aspects of mathematics such as proofs of theorems, or graphing. An instructor may feel that all students should be helped with all problems. But with limited resources, what kind of help should be given? In the final analysis it is up to the professor to set such priorities, and to choose those resources which best fit the needs of the student in his or her particular class, within the constraints of limited resources.

Conclusion. Different tutor resources have different properties and appeal to different students. It is of course an immense (and probably impossible) task to gather enough data to be able to predict student preferences, although this may be of interest to educators. The best we as teachers can do is to make available to students a sufficient variety of resources and not require of every resource that it be popular with a large majority of students.

CONSULTING GROUP ON INSTRUCTIONAL DESIGN, 205 ELLIOTT HALL, UNIVERSITY OF MINNESOTA, MINNEAPOLIS, MN 55455.

REMEDIAL MATHEMATICS: AN ADMINISTRATOR'S VIEWPOINT

MICHAEL C. GEMIGNANI

IUPUI (Indiana University–Purdue University–Indianapolis) is a large state university located in a sizable metropolitan area. While not having an "open admission" policy, we find a rapidly increasing number of students seeking remedial courses (equivalent to first and second year high school algebra). Whereas one or two sections a semester of MATH 001 (first year high school algebra, no academic credit) sufficed in 1972–73, this past academic year six sections a semester could barely accommodate the demand. Since funds available to create new sections are limited, steps had to be taken to accommodate the additional students without a disruptive increase in cost. The principal action taken was to implement larger sections of MATH 111 (second year high school algebra, degree credit only in a few programs). To try to determine what effect, if any, the increase in class size had on the students' performance, I decided to compare Spring 72–73, the first semester for which readily available records exist, and Spring 75–76, the last full semester prior to the study.

In 1975–76 various aids were made available to students in remedial mathematics. These included free tutorial sessions (about 20 hours per week), which were also available on a more limited basis during 72–73, modules of various lessons on tape in the library, and a testing center to permit students in MATH 001 and 111 to make up tests on which they had done poorly, or for which they had been absent.

Several changes were also made in the content of MATH 111, but the same texts were used in 75–76 as in 72–73. Any comparison of the results in the two semesters is complicated by certain grading policies implemented in the interim. For example, it was decided not to award D's in MATH 111 in

TABLE 1. *Enrollment for MATH 001*

	number of students	number of sections	average size of a section
Spring 72–73	56	1	56
Spring 75–76	322	6	53.67

TABLE 2. *Enrollment for MATH 111*

	number of students	number of sections	average size of a section
Spring 72–73	231	5	46.2
Spring 75–76	497	5	99.4

75–76, but instructors were generally more lenient in giving some students C's who might otherwise have received D's. The most significant difference between the two semesters was the number of students and the average class size in MATH 111. Was the rate of success in MATH 111 in 72–73 significantly better than that of 75–76?

TABLE 3. *Grade distribution in MATH 111 by %**

	Spring 72–73 (%)	Spring 75–76 (%)
A	14.7	8.8
B	16.44	10.81
C	18.98	9.37
D	7.82	0
F	13.24	21.38
Withdrew	28.8	28.92
Credit gained	58.94	44.74
No credit	42.04	50.3

*Failure to sum to 100 due to rounding error or students receiving incomplete grade report.

If the success-fail rate had been the same in 75–76 as in 72–73, then 293 students would have received credit for MATH 111, but only 222 actually did; thus it appears that 71 additional students would have passed MATH 111 in 75–76 had we stayed with the average class size we had in 72–73. Using Chi-square, the data above appears to confirm the hypothesis that larger class size means higher failure rate, and this at a .001 level of significance.

My use of statistics, though, to try to prove anything from the data available is, of course, quite out of order for a number of reasons even though there does definitely seem to be a higher failure rate in MATH 111 in 75–76. It is also unconscionable to blame this failure rate in whole or in part on the larger sections without additional analysis, since 75–76 may be different from 72–73 in other important respects as well. For example, are the instructors better in 72–73 (my knowledge of the instructors tells me no)? Is the course harder (some changes may have made it so)? Does the exclusion of the D appreciably affect matters? Perhaps. Are the students worse now than they were in 72–73?

There are indeed data to support the hypothesis that the students were of lower quality in 75–76 than in 72–73. In Table 4 we list the expected values for students passing and failing MATH 001 assuming that the pass-fail rate in 75–76 is the same as that of 72–73. Below those we list the actually observed values. Using our suspect Chi-square again, we have support for a hypothesis, this time at a 5% level of significance. The students in MATH 001 in 75–76 are not as capable as those in 72–73.

The class size in MATH 001 remained about the same from 72–73 to 75–76, yet the success rate fell from 42.8% to 37.37%. Note too that in only one semester in only one course (MATH 111 in 72–73) was the success rate higher than 50%, and in MATH 001 in 75–76 just about one student in three who

TABLE 4. *MATH* 001 75–76 *performance*

	pass	fail
# expected using 72–73 rate	139	183
actual # observed	120	202

registered for the course actually completed it with a passing grade. Are those who failed in previous semesters coming back again and again to repeat these courses, thus forcing the enrollment to rise even more rapidly each year? Are we doing something dramatically wrong with our instruction? Are we doing the job that the State of Indiana tells us we should by entrusting public money to our care?

But here is where my administrator's conscience furnishes a counterpoint against my instructor's compassion for all those students who are trying, but not making it in MATH 001 and 111. Let us suppose, for example, that class size is a critical element in the difference between the success rates in the two semesters. We would then have salvaged an additional 14.2% of the students, or approximately 71 souls, had we preserved the average section size of 46.2. But elementary arithmetic tells us that we would have required 11 sections of MATH 111 in 75–76 instead of 5, an increase of 6 sections. Every member of the academic community who aspires to administration knows that you don't get something for nothing, except trouble; therefore we would have had to spend approximately twice as much to teach MATH 111 for those additional 71 students to pass as we did to provide an environment in which they failed. Put another way, since about three times 71 students passed anyway, we would have had to double our expense to increase the number of successful students by 1/3.

The highest rate of success anyone has ever quoted to me about a course like MATH 001 or 111 is 66% (please let me know if you have done better and how you did it) and that by giving the students almost individual attention. By reducing the class size to 10, let us say, we could have theoretically saved an additional few percent, about 1/12, at a four-fold increase in expense. The Law of Diminishing Return clearly begins to operate beyond a certain point.

Now I can hear many crying, "A logical extension of your argument is that we really shouldn't spend anything at all on them. Why not make each student in MATH 001 and 111 buy a book and report to an examination at the end of the semester. A few will pass, most will fail, but look at all the money saved."

A seemingly flippant and insensitive response is that perhaps that is the way it should be done. But the response has an element of truth in it.

A former head of Business at IUPUI once remarked to me that he had never known a student to successfully complete the four year program in his school who had to start by taking MATH 001. His statement is probably but a slight exaggeration. Any student who begins college having had no high school algebra whatsoever is in pretty bad shape. He could be a bright student who was miscounseled or unmotivated in high school, but who has now found his bearings and wants to push ahead. Such a student will probably do well in remedial courses no matter how they are presented. Or he could be a student who shouldn't be in college to begin with. If this student is gently eased through what he should have had in his early years of high school only to flunk out once he reaches demanding upper division courses, are we not misleading him in the first place? Are we justified in taking his money to teach him what he should have learned before coming to us, only to flunk him when he reaches truly college level work?

I can appreciate the argument that it is better that persons needing remedial mathematics upgrade their skills even when they find they cannot go on to more advanced work; the more educated a person is, the better, even if he isn't "fully" educated. I can also sympathize with instructors and students who find it burdensome to teach, or attend, a large lecture section, even with teaching assistants to lighten the load and provide greater personal contact. Yet it becomes in the end, at least

in my mind, a question of where the limited funds available are to be spent. Where do they give the most "return"? And, indeed, here I find myself in a morass of philosophical and social questions which have almost as many answers as persons attempting answers.

My own view is that funds for colleges and universities should be spent primarily on instruction at a level appropriate to a college or university. While society and our private consciences may insist that we give the educationally disadvantaged an opportunity to prepare themselves for college level work, even while they are actually attending a college, I do not feel that it does either the student, the institution, or even society, a favor by diverting funds intended for higher education to this purpose.

An address to the Indiana Section of the MAA, November 1976.

DEPARTMENT OF MATHEMATICAL SCIENCES, INDIANA UNIVERSITY-PURDUE UNIVERSITY AT INDIANAPOLIS, INDIANAPOLIS, IN 46205.

AN EXPERIMENTAL MATHEMATICS PROJECT FOR WOMEN

CAROLYN T. MACDONALD AND BARBARA S. CURRIER

Many careers, particularly those in the sciences, traditionally have been filled by men. There are many complex reasons why this is the case, but one major factor is that women often have not been encouraged to pursue sufficient mathematics during their high school years. For example, in a fact sheet on women in higher education at the University of California–Berkeley, sociologist Lucy Sells reported that in a random sample of freshmen admitted in Fall, 1972, 57% of the males had completed four full years of high school mathematics as compared with only 8% of the women. Such disparities, which can be found at most schools, mean that the deficient mathematical backgrounds of women have precluded their considering many career options, in the sciences as well as in a great many other areas.

Project. This situation was the motivation for an experimental project conducted at the University of Missouri–Kansas City during the 1974–75 academic year. The project, which was funded by a grant from the National Science Foundation, was designed to open the career options of women students by strengthening their mathematical backgrounds and, just as important, to assist them in overcoming the social, cultural, and psychological barriers to their success.

The project consisted of offering a special section of the introductory mathematics sequence, restricted to women students. Math 110, which was offered in the Fall, 1974, is the approximate equivalent of college algebra, including probability; and Math 120, which was offered in the Spring, 1975, is the equivalent of trigonometry and analytic geometry. Math 110 is taken by most students with less than four years of high school mathematics as part of a basic skills graduation requirement.

The experimental section differed from the standard sections of Math 110 and Math 120 in a number of ways. During the first semester the experimental section consisted of thirty-two women of varying ages and backgrounds who were selected for participation because they lacked a strong preparation in mathematics. The experimental section was team-taught by Carolyn MacDonald, the project director and director of the Physical Science Program at UMKC, and Barbara Currier, an advanced doctoral student in mathematics, with both instructors attending all of the daily class sessions. In contrast, standard sections average 55–60 students, with about half women, and are taught by a single instructor who may be a regular faculty member, non-regular instructor, or graduate assistant.

There are regular tutoring sessions staffed by graduate assistants which are open several hours a day to students in all Math 110 sections. For the experimental section, this was replaced by a daily tutoring session staffed by an undergraduate held for participants only during the hour immediately preceding class.

The text used in the experimental section was oriented more towards problem-solving, containing more examples and exercises than did the text used in the standard sections. Because of the weak backgrounds of many of the participants, the first four weeks of the first semester were spent on a review of arithmetic, integrated with instruction on use of the slide rule.

Results. In all sections of Math 110 grades are assigned according to a standardized percentage scale, 90–100%, A, etc. This usually results in final grades of about one fifth each of A, B, C, combined D and F, and Withdraw. In the experimental section the grade distribution was 53%, A; 25%, B; 6%, C; 9%, D; 3%, F; and 3%, Withdraw. Dr. MacDonald had taught two different sections of Math 110 during the previous two years, one a standard section and one a section for students with weak backgrounds comparable to those of the program participants. In those sections her final grades were similar to the usual grade distribution. In her judgment, however, the exams given in those sections were less demanding than those given the experimental section. Thus the higher grades should not be attributed to easier exams or grading.

Although the first four weeks of the semester were spent on basic arithmetic review, the experimental section actually covered considerably more material by the end of the semester than is generally taught in Math 110. In fact, it was largely due to the time spent on review that this was possible, because the students then had a stronger background and greater confidence in their ability to learn mathematics.

The lower rate of withdrawal from the first semester and the higher rate of continuation to a second semester were also indications of the success of the project. Only one out of 32 participants withdrew from the first semester, versus 22% of women in the standard sections. Thirty-eight percent of the students in the project went on to complete with a grade of C or better a mathematics course the following semester. This rate of continuation exceeded somewhat the normal rate of continuation for men and far exceeded the usual rate for women with comparable backgrounds. In classes previously taught by Dr. MacDonald, in the standard section 22% of the men and 24% of the women continued, while in the section for students with weak backgrounds, 32% of the men and only 3% of the women completed a mathematics course the next semester.

Another measure of the effectiveness of the program is the change in participants' attitudes about mathematics. In a computer-scored questionnaire administered to all sections of Math 110 at the time of the final exam, 76% of the participants as compared to 40% of the non-participant women and 47% of the men reported that their current understanding of mathematics was much better, and 36% of the participants as compared to 17% of the non-participant women and 15% of the men reported that their interest in mathematics was much higher than before they enrolled in Math 110. At least four of the participating women have altered their career plans as a result of the project, changing to majors that require more mathematics and traditionally have few women.

Analysis. All of the differences between the experimental section and the standard sections may have contributed to the success of the project. Most of the participants felt that a major factor was the absence of competition and intimidation that they might have felt if there had been male students in the class. The class developed a unique atmosphere of friendship and cooperation which they believed would have been impossible in a mixed group. The development of class unity also was aided by the special tutoring section, which took on a definite social aspect in addition to the intended functions.

The fact that the class was team-taught was more an aid to the instructors than it was to the students, although it did result in more time available for the instructors to help students on an individual basis.

Conclusion. This project has shown that it is possible to break down the barriers against mathematics that many students, particularly women, bring with them to college. Many of the factors which develop the necessary supportive environment are inexpensive from a financial standpoint, although they may be costly in terms of time and devotion.

Note. Interested readers may obtain a more comprehensive report by writing Dr. MacDonald.

This project was supported by a National Science Foundation grant, GY 11326.

References

1. J. Ernest and a research team, Mathematics and Sex, Mathematics Department, University of California–Santa Barbara, January, 1975.
2. C. MacDonald, Increasing Women in the Sciences through an Experimental Mathematics Project, Final report, October, 1975.
3. L. Sells, High School Mathematics as the Critical Filter in the Job Market, Fact sheet, March 31, 1973.

MATHEMATICS DEPARTMENT, BAPTIST COLLEGE, CHARLESTON, SC 29411.
MATHEMATICS DEPARTMENT, ROCKHURST COLLEGE, AND MATHEMATICS DEPARTMENT, UNIVERSITY OF MISSOURI–KANSAS CITY, KANSAS CITY, MO 64110.

PROBLEMS AND SOLUTIONS

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before September 30, 1977.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

E 2659. *Proposed by Arthur L. Holshouser, Charlotte, North Carolina*

The sequence a, b, c, d can be parenthesized in 5 ways. Equating these two at a time we obtain the following "identities":

$$(1)\ (ab)(cd)=a(b(cd)),\qquad (2)\ (ab)(cd)=a((bc)d),$$
$$(3)\ (ab)(cd)=((ab)c)d,\qquad (4)\ (ab)(cd)=(a(bc))d,$$
$$(5)\ (a(b(cd))=a((bc)d),\qquad (6)\ a(b(cd))=((ab)c)d,$$
$$(7)\ a((bc)d)=(a(bc))d,\qquad (8)\ ((ab)c)d=(a(bc))d,$$
$$(9)\ a(b(cd))=(a(bc))d,\qquad (10)\ a((bc)d)=((ab)c)d.$$

Which of these identities implies that a quasigroup satisfying it is necessarily a group?

E 2660. *Proposed by E. Ehrhart, Strasbourg, France*

A quadrilateral is *cyclic* if its vertices lie on a circle. Find the number of congruence classes of cyclic quadrilaterals having integral sides and given perimeter n.

E 2661. *Proposed by Steve Galovich, Carleton College, Northfield, Minnesota*

Find all functions f which satisfy the three conditions

$$\text{(i) } f(x,x)=x; \qquad \text{(ii) } f(x,y)=f(y,x);$$
$$\text{(iii) } (x+y)f(x,y)=yf(x,x+y),$$

assuming that the variables and the values of f are positive integers.

E 2662.* *Proposed by Edward T. H. Wang, Wilfrid Laurier University, Ontario*

For an $n \times n$ (0,1)-matrix A, let A' denote the complementary matrix, i.e., $A' = J - A$ where J is the matrix with all entries equal to one. Define $\sigma_n = \max \Sigma(AA')$ where $\Sigma(X)$ denotes the sum of all entries of a matrix X and the maximum is taken over all $n \times n$ (0,1)-matrices A.

Show that $\sigma_n \geqq (n^3 - n)/3$.* Does the equality hold for all n?

E 2663. *Proposed by Marius Solomon, Student, University of Pennsylvania*

Let $f:(0,\infty)\to\mathbf{R}$ be differentiable and assume that $f(x)+f'(x)\to 0$, when $x\to\infty$. Show that $f(x)\to 0$ as $x\to\infty$.

E 2664. *Proposed by Robert L. Bishop, Massachusetts Institute of Technology*

For a fixed $n \geqq 3$ describe how one can construct all solutions of the system of Diophantine equations

$$\left(\sum_{i=1}^{n} x_i\right) - x_j = y_j^2, \qquad 1 \leqq j \leqq n.$$

For $n = 10$ find a solution such that the x_i are distinct positive integers and $x_1 + \cdots + x_{10}$ is minimal.

SOLUTIONS OF ELEMENTARY PROBLEMS

A Binary Operation in the Plane

E 2579 [1976, 133]. *Proposed by Benjamin Klein and Brian White, Davidson College, North Carolina*

Let $0 < \theta < \frac{1}{2}\pi$ and let p, q be arbitrary distinct points in the Euclidean plane E. Define $f_\theta(p,q)$ to be the unique point r in E such that triangle pqr is in the counterclockwise sense and $\measuredangle rpq = \measuredangle rqp = \theta$ radians. Show that $f_{\pi/3}(p,q)$ can be written as an expression involving only $f_{\pi/6}$, p, q, and parentheses.

Solution by the solvers listed below. Let

$$g(p,q) = f_{\pi/6}(p, f_{\pi/6}(p, f_{\pi/6}(p,q))).$$

Then it is easy to check that p, q, $g(q,p)$, $g(p,q)$ is a rectangle (in the counterclockwise sense) and if the distance between p and q is 1 then the distance between p and $g(p,q)$ is $1/3\sqrt{3}$.

Define functions g' and g'' by

$$g'(p,q) = g(g(p,q), g(q,p)), \quad g''(p,q) = g(g'(p,q), g'(q,p)).$$

Then we have $f_{\pi/3}(p,q) = f_{\pi/6}(g''(p,q), g''(q,p))$.

Solved by Anders Bager (Denmark), Peter de Buda, Landy Godbold, L. E. Mattics, J. G. Mauldon, Hugh Noland, James Ridley (South Africa), and the proposers.

Chebyshev Polynomials

E 2580 [1976, 133]. *Proposed by Clark Kimberling, University of Evansville*

Show that

$$\frac{d}{dx}\left[\prod_{k=0}^{n-1}\left(a - 2\sqrt{x}\cos\frac{(2k+1)\pi}{2n}\right)\right] = -n\prod_{k=1}^{n-2}\left(a - 2\sqrt{x}\cos\frac{k\pi}{n-1}\right).$$

Solution by A. A. Jagers, Technische Hogeschool Twente, Enschede, Netherlands. The Chebyshev polynomials T_n and U_{n-2} are given by

$$\cos n\theta = T_n(\cos\theta) = 2^{n-1}\prod_{k=0}^{n-1}\left(\cos\theta - \cos\frac{(2k+1)\pi}{n}\right), \tag{1}$$

$$\frac{\sin(n-1)\theta}{\sin\theta} = U_{n-2}(\cos\theta) = 2^{n-2}\prod_{k=1}^{n-2}\left(\cos\theta - \cos\frac{k\pi}{n-1}\right). \tag{2}$$

Thus

$$\prod_{k=0}^{n-1}\left(a - 2\sqrt{x}\cos\frac{(2k+1)\pi}{2n}\right) = 2(\sqrt{x})^n\, T_n\left(\frac{a}{2\sqrt{x}}\right),$$

$$\prod_{k=1}^{n-2}\left(a - 2\sqrt{x}\cos\frac{k\pi}{n-1}\right) = (\sqrt{x})^{n-2}\, U_{n-2}\left(\frac{a}{2\sqrt{x}}\right).$$

Put $z = a/2\sqrt{x}$. Then the identity to be proved may be written

$$nT_n(z) - zT_n'(z) = -nU_{n-2}(z). \tag{3}$$

But (3) follows from (1), (2) and $\sin(n-1)\theta = \sin n\theta\cos\theta - \cos n\theta\sin\theta$.

Also solved by Anders Bager (Denmark), Clark Givens, I. I. Kolodner, L. E. Mattics, T. M. Mills (Australia), Ram Murty & Kumar Murty (Canada), Otto Ruehr, St. Olaf Problem Group, David Zeitlin, and the proposer.

A Property of Fibonacci Numbers

E 2581 [1976, 197]. *Proposed by Clark Kimberling, University of Evansville*

Let ϕ denote Euler's totient function and let $\{F_n\}$ be the sequence of Fibonacci numbers: $F_1 = F_2 = 1$ and $F_{n+2} = F_n + F_{n+1}$. Show that $\phi(F_n)$ is divisible by 4 if $n \geqq 5$.

Solution by Peter L. Montgomery, Huntsville, Alabama. The four integers ± 1, $\pm F_{n-1}$ are non-congruent modulo F_n when $n \geqq 5$. It is well known that $F_{n-1}^2 - F_{n-2}F_n = (-1)^n$ and consequently the above four integers form a subgroup of the multiplicative group G of units of the ring of integers modulo F_n. Since $|G| = \phi(F_n)$ the assertion follows from Lagrange's Theorem.

Also solved by Anders Bager (Denmark), George Berzsenyi, D. M. Bloom, Richard Bauer, Robert Breusch, Peter de Buda, José Luis de Miguel (Spain), Ira Gessel, Richard Gibbs, Mark Goldsmith, Bill Heidler, Verner

Hoggatt & L. Kuipers (Switzerland), A. A. Jagers (Netherlands), Victor Keiser, S. C. Locke (Canada), Graham Lord (Canada), Helen Marston, L. E. Mattics, Ram Murty & Kumar Murty (Canada), D. E. Penney, Ann Playtis, Bob Prielipp, Ira Rosenholtz, Sahib Singh, Edith Sloan, Lawrence Somer, E. Trost (Switzerland), G. W. Valk, and the proposer.

Editor's Comment. Prielipp notes that this problem was proposed in the Fibonacci Quarterly with solution published in December 1966, pp. 334–335.

Crisscrossing Partitions of a Finite Set

E 2582 [1976, 197]. *Proposed by Ioan Tomescu, University of Bucharest, Rumania*

Let $\{A_i: 1 \leqq i \leqq n\}$, $\{B_i: 1 \leqq i \leqq n\}$ and $\{C_i: 1 \leqq i \leqq n\}$ be three partitions of a finite set M. If for every i, j, k we have

$$|A_i \cap B_j| + |A_i \cap C_k| + |B_j \cap C_k| \geqq n,$$

prove that $|M| \geqq n^3/3$, and that this inequality cannot be improved when n is divisible by 3.

Composite of (independent) solutions by Bernhardt Ganter, Technische Hochschule Darmstadt, Germany, and Dwight R. Bean, University of San Diego. We have

$$\sum_{i,j} |A_i \cap B_j| = \sum_{i,k} |A_i \cap C_k| = \sum_{j,k} |B_j \cap C_k| = |M|.$$

From the hypothesis,

$$\sum_{i,j,k} (|A_i \cap B_j| + |A_i \cap C_k| + |B_j \cap C_k|) \geqq \sum_{i,j,k} n = n^4.$$

Thus $3n|M| \geqq n^4$, and $|M| \geqq n^3/3$.

Now consider $M = \{(i, j, k): 1 \leqq i, j, k \leqq n \text{ and } i + j + k \equiv 0 \pmod 3\}$. Let A_i be the set of points in M with first coordinate equal to i. Similarly B_j consists of points with second coordinate j and C_k of points with third coordinate k. Then, in case n is a multiple of 3, it is easy to check that $|A_i \cap B_j| = |A_i \cap C_k| = |B_j \cap C_k| = n/3$ and that $M = n^3/3$.

Also solved by Anders Bager (Denmark), Peter de Buda, Michael Josephy (Costa Rica), Dana Mabbott, Peter Montgomery, Robert Patenaude, Adam Riese, Eric Rosenthal, Joel Spencer, University of South Alabama Problem Group, James Walker, Philip Washburn, and the proposer. Partial solution by David Bienenfeld (Israel), and José Luis de Miguel (Spain).

Infinite 2-complex in 3-space

E 2584 [1976, 198]. *Proposed by H. S. M. Coxeter, University of Toronto*

Describe an infinite complex of congruent isosceles triangles, extending systematically throughout three-dimensional Euclidean space in such a way that each side of every triangle belongs to just two other triangles.

I. Example by Carl Pomerance, University of Georgia, and Rodney T. Hood, Franklin College, Indiana (independently). Consider the standard decomposition of $\mathbf{R}^3$ into unit cubes centered in the lattice points $\mathbf{Z}^3$. Each lattice point $(a, b, c) \in \mathbf{Z}^3$ together with each of the 12 edges of the cube $(a \pm \frac{1}{2}, b \pm \frac{1}{2}, c \pm \frac{1}{2})$ determines 12 isosceles triangles. Let $C(a, b, c)$ be the finite complex consisting of these 12 triangles. The union of $C(a, b, c)$ taken over all $(a, b, c) \in \mathbf{Z}^3$ such that a, b, c are not all even or all odd, has the required properties.

II. *Example by Clarence R. Perisho, Mankato State University, Michigan (revised by the editor).* Let Γ_0 be a complex consisting of 6 triangles each of which has the origin as one vertex and the remaining

two vertices are, respectively

(1) $(\pm 1, 1, 1)$ (2) $(\pm 1, -1, 1)$

(3) $(2,0,0), (1,1,1)$ (4) $(2,0,0), (1,-1,1)$

(5) $(1, \pm 1, 1)$ (6) $(0,2,0), (-1,1,1)$.

Let $\tau_{\mathbf{v}}$ be the translation of $\mathbf{R}^3$ by the vector $\mathbf{v}$, and ρ be the rotation for $90°$ about the z-axis. Then

$$\Gamma' = \bigcup_{a,b \in \mathbf{Z}} \tau_{(2a,2b,0)} \Gamma_0$$

is an infinite complex. It has the following properties:

(a) Every non-horizontal edge of Γ' belongs to exactly 3 triangles.

(b) Every edge of Γ' which is parallel to the x-axis (resp. y-axis) belongs to exactly 2 triangles (resp. 1 triangle).

It suffices to check (a) and (b) only for edges that belong to Γ_0, which is easy.

Now let

$$\Gamma'' = \bigcup_{c \in \mathbf{Z}} \tau_{(0,0,2c)} \Gamma'$$

and

$$\Gamma = \Gamma'' \cup \tau_{(1,1,1)} \rho \Gamma'.$$

It follows from (a) and (b) that the complex Γ has the required properties. Moreover, Γ retains these properties after a stretching in the direction of the z-axis.

III. *Example by the proposer and Richard S. Stevens (independently)*. Consider the decomposition of 3-space into congruent truncated octahedra (see Hugo Steinhaus, *Mathematical Snapshots*, 2nd ed., Oxford University Press, New York 1950, p. 156). The proposed complex consists of all the triangles that join the center of each cell to its 36 edges.

Average Vertex-Degree for Triangulated Surfaces

E 2585 [1976, 198]. *Proposed by Jan Mycielski, University of Colorado*

Prove that for every triangulation of a 2-dimensional closed surface, the average number of edges meeting at a vertex approaches 6 in the limit as the number of triangles used approaches infinity.

Solution by Lee Erlebach, Michigan Technological University. Denote the Euler characteristic of the surface by χ and denote by v, e, f the number of vertices, edges and faces, respectively, in a triangulation of the surface. Then by Euler's theorem, $v - e + f = \chi$. By counting in two ways the ordered pairs (T, x) where T is a face and x a vertex of T, we obtain $2e = 3f$. For any triangulation, we have that the average number d of edges meeting at a vertex is $d = 2e/v$. Using the above equations we get $d = 6f/(f + 2\chi)$. Thus $d \to 6$ as $f \to \infty$. Note that $d = 6$ if $\chi = 0$, i.e., if our surface is a torus or a Klein bottle.

Also solved by Anthony Barkauskas, Dwight Bean, H. E. Bible Jr., & S. R. Murdock, Irl Bivins, Ben Burrell, Paul Chernoff, Richard Gibbs, Rodney Hood, Lawrence House, Morris Marx, Robert Patenaude, D. E. Penney, Ken Rebman, Adam Riese, Eric Rosenthal, Robin Soloway, Joel Spencer, Richard Stevens, Warren White, and the proposer.

Chernoff notes that if n-gons are used instead of triangles then $d \to 2n/(n-2)$ as $f \to \infty$.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Hill Center, Rutgers University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate signed sheets and should be mailed before September 30, 1977.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

6156. *Proposed by Herbert Knothe, Bremen, Germany*

Prove: If a prime p has the form $8n+7$ then the number of even quadratic residues $>p/2$ equals $n+1$. If a prime p has the form $8n+3$ then the number of even quadratic residues $<p/2$ equals n. (Each residue r is restricted so that $0 \leqq r < p-1$.)

6157*. *Proposed by C. C. Chen, Nanyang University, Singapore, and D. E. Daykin, Reading University, England*

(a) Find integers Δ, p with the following property: Whenever the lines of the complete graph K_p are colored so that every vertex is on $\leqq \Delta$ lines of each color, there is a triangle whose lines have different colors.

(b) Find integers δ, p, n with the following property: Whenever the lines of a complete graph K_p are colored with n colors so that every vertex is on $\geqq \delta$ lines of each color, there is a triangle whose lines have different colors.

6158*. *Proposed by M. J. Pelling, University of Benin, Nigeria*

Prove that if R is a bounded convex region of the plane of area 1 then there is a $d>0$ independent of R such that R is equivalent under an area preserving affine transformation to a region of diameter $\leqq d$. What is the best possible value of d?

6159. *Proposed by Thomas E. Elsner, General Motors Institute*

It is well known that for a graph on k vertices with no triangles the maximum number of edges is $L(k)=mn$ where $m=[k/2]$ and $n=[(k+1)/2]$ and that this value occurs for the complete bigraph $K_{m,n}$. Express the maximum number of edges in case we add the restriction that the graph be (a) Hamiltonian; (b) Eulerian.

6160. *Proposed by Robert E. Shafer, Berkeley, California*

(1) If m is the largest odd divisor of n, then with the exception of (2),

$$2^{\nu(n)}m^{\nu(m)/2} \mid \phi(a^n+b^n)$$

for $a>b\geqq 1$, $\nu(n)$ is the number of divisors of n, ϕ is the Euler function.

(2) If $a=2$, $b=1$, $n=3^dc$, c odd, $d \geqq 1$,

$$2^{\nu(n)-1}m^{\nu(m)/2}3^{d-1} \mid \phi(a^n+b^n).$$

6161. *Proposed by Clark Kimberling, University of Evansville*

For $0<r<1$, let $S(r)$ be the set of integers n such that one and only one integer lies in the open interval $(nr, nr+r)$. Prove or disprove that r is irrational if and only if, for every positive integer M, the set $S(r)$ contains a complete residue system modulo M.

SOLUTIONS OF ADVANCED PROBLEMS

Extending a Sublinear Map

6051* [1975, 857]. *Proposed by Jochem Zowe, University of Würzburg, Germany*

Let X be a real vector space, Y an ordered vector space and p a sublinear map of X into Y, i.e., $p(\lambda x) = \lambda p(x)$ and $p(x + x') \leqq p(x) + p(x')$ for all $x, x' \in X$ and all real nonnegative λ. Does there always exist a linear map T of X into Y such that $Tx \leqq p(x)$ for all $x \in X$?

Partial solution by Anthony G. O'Farrell, St. Patrick's College, Kildare, Ireland. Let X be a real vector space, Y an *order-complete* ordered vector space, and P a sublinear map of X into Y, then there exists a linear map $T: X \to Y$ such that $Tx \leqq p(x)$ for all x in X.

Proof. Consider the set $\mathcal{F}$ of pairs (U, S), where U is a subspace of X and $S: U \to Y$ is a linear map with $Su \leqq p(u)$ for all $u \in U$. Partially order $\mathcal{F}$ by $\leqq$, where $(U, S) \leqq (U_1, S_1)$ means $U \subseteq U_1$ and the restriction of S_1 to U is S.

Since $p(0) \geqq 0$, the pair $((0), 0)$ belongs to $\mathcal{F}$, hence $\mathcal{F} \neq \emptyset$. Zorn's lemma applies and shows that there exists a maximal (U, S). If $U \neq X$, choose $\xi \in X - U$. For positive α and β in $\mathbb{R}$ and $u, v \in U$, we have

$$S(\alpha v + \beta u) \leqq p(\alpha v + \beta u) \leqq p(\alpha v - \alpha\beta\xi) + p(\beta u + \alpha\beta\xi) = \alpha p(v - \beta\xi) + \beta p(u + \alpha\xi),$$

$$\frac{1}{\beta}\{Sv - p(v - \beta\xi)\} \leqq \frac{1}{\alpha}\{p(u + \alpha\xi) - Su\}.$$

Since Y is order-complete,

$$m = \text{l.u.b.}\ \frac{1}{\beta}\{Sv - p(v - \beta\xi)\}$$

exists, where β ranges over all positive numbers, and v ranges over U. For $\alpha > 0$ in $\mathbb{R}$ and $u \in U$ we have

$$(*) \qquad Su - p(u - \alpha\xi) \leqq \alpha m \leqq p(u + \alpha\xi) - Su.$$

Define S_1 on $U + \mathbb{R}\xi$ by $S_1(u + \gamma\xi) = Su + \gamma m$. Then we have by $(*)$, checking alternately $\gamma > 0$, $\gamma < 0$,

$$S_1(u + \gamma\xi) \leqq p(u + \gamma\xi),$$

$\gamma \in \mathbb{R}$ and $u \in U$. This contradicts the maximality of (U, S). Thus $U = X$, and $T = S$ is the desired map.

Also solved in the case of an order complete range by Michael Barr, and by the proposer. For general Y the problem remains open.

Torsion Groups Generated by Two Elements

6052 [1975, 857]. *Proposed by J. R. Gard, University of South Florida*

If G is a torsion group such that there exists an element $x \in G$ with the property that x and y generate G whenever $y \in G$ is not a power of x, is G finite? What other properties does G have?

Solution by S. G. Udpikar, S. P. College, Poona, India. A group G with the given property is finite if and only if it is solvable.

Assume that G is finite. Let $M = (x)$. This is a finite abelian subgroup of G. Also it is a maximal subgroup of G. For suppose that $M \subset M^*$ and $M \neq M^*$. Then there exists $y \in M^*$ such that $y \notin M$, that is, y is not a power of x. Now $x, y \in M^*$. Hence $M^* = G$. Moreover, M is abelian. Hence, [1] Chapter 7, $G''' = \{e\}$. Hence G is solvable.

Conversely, assume that G is solvable. As G is a periodic solvable group, it is, [2] Chapter VII, locally finite.

Case (i): Let G be abelian. Then $G' = \{e\}$. In this case $M = (x)$ is a normal subgroup of G. Let $y \in G$ be such that $y \notin M$. (If there exists no such $y \in G$, then $M = G$ and hence G is finite.) Let $H = (y)$. Consider the subgroup MH of G. $y \in MH$, but $y \notin M$. Therefore, $MH \supset M$ properly. But M is a maximal subgroup of G, therefore $MH = G$. But $O(M)$, $O(H) < \infty$. Hence $O(G) = O(M)O(H)/O(M \cap H) < \infty$.

Case (ii): Let G be non-abelian, so that $G' \neq \{e\}$. Now if $M \supset G'$, then M is normal in G. Hence G is finite as in case (i) above. Assume that M does not contain G' properly. Therefore there exists $y \in G'$ such that $y \notin M$. Now, as before, $G = MG'$. But $G/G' \cong MG'/G' \cong M/M \cap G'$, which is finite. Therefore G' is a subgroup of finite index in a finitely generated group G. Hence G' is finitely generated. See [2] Chap. V. Hence G' is finite as G is locally finite. Now, as $G = MG'$, G is finite (as above).

References

1. John D. Dixon, Problems in Group Theory, Blaisdell (1967).
2. E. Schenkman, Group Theory, Van Nostrand (1965).

Density of Arguments of Powers of Gaussian Integers

6053 [1975, 857]. *Proposed by Raphael Finkelstein, Bowling Green State University*

Let $a + bi$ be a Gaussian integer with $(a, b) = 1$ and let $A + Bi = (a + bi)^p$ where p is an odd prime. Let $C = \max(A, B)$ and $D = \min(A, B)$. Can C/D approach $\frac{1}{2}(1 + \sqrt{5})$ arbitrarily closely?

Solution by J. C. Lagarias, Bell Laboratories, Murray Hill, New Jersey. Let $\alpha = A + Bi$ be a fixed element of $Z[i]$ given in advance, and let $\alpha^p = A_p + B_p i$. We claim that if α is not an integral multiple of 1, i, $1 + i$ or $1 - i$, then the set

$$S(\alpha) = \{B_p/A_p \mid p \text{ an odd prime}, A_p \neq 0\}$$

is dense in $\mathbb{R}$, and thus in particular, includes $\frac{1}{2}(1 + \sqrt{5})$ as a limit point.

To show this let $\alpha = re^{i\theta_0}$ with θ_0 defined (mod 2π). Now $\alpha^p = r^p e^{ip\theta_0}$ and

$$B_p/A_p = \tan p\theta_0.$$

Since $\tan\theta\colon(-\frac{1}{2}\pi, \frac{1}{2}\pi) \to \mathbb{R}$ is continuous and onto, it is sufficient to show

$$S_{\theta_0} = \{p\theta_0 (\text{mod } 2\pi) \mid p \text{ prime}\}$$

is dense in $[0, 2\pi)$ for the given θ_0. For a general θ, the behavior of S_θ is described by:

(i) S_θ is a finite set if $\theta/2\pi$ is rational

(ii) S_θ is dense (in fact, uniformly distributed) in $[0, 2\pi)$ if $\theta/2\pi$ is irrational.

This follows from a theorem of I. M. Vinogradov [*The Method of Trigonometrical Sums in the Theory of Numbers*, (Interscience), notes on p. 120]. It remains to determine which α have argument θ_0 with $\theta_0/2\pi$ irrational, and which have $\theta_0/2\pi$ rational. Now $\tan\theta_0 = B/A$ is rational, so the appropriate question to ask is: "For which rational r is $\tan 2\pi r$ rational?" The answer is that r must be an integral multiple of $\frac{1}{4}$ (Corollary 3.12 in I. Niven, *Irrational Numbers*, Carus Mathematical Monographs # 13.) Thus we are led to exclude $\alpha = 1, i, 1 + i, 1 - i$; for all other α in $Z[i]$, $\theta_0/2\pi$ is irrational and (ii) holds.

Also solved by L. Kuipers (Switzerland).

Note. The proposer offered the following comment in submitting the problem: this problem, if the answer had been negative, would yield the complete solution of all the pth powers in the Fibonacci sequence.

Mapping Induced by a Permutation

6054 [1975, 941]. *Proposed by Lung Ock Chung, University of California at Los Angeles*

Let $\varphi:\{0,1,\ldots,N-1\}\to\{0,1,\ldots,N-1\}$ be a permutation for $N\geqq 2$. Then φ induces a function $\varphi^*:(0,1)\to(0,1)$ from the open unit interval to itself as

$$\varphi^*\left(\sum_{i=1}^{\infty}\frac{m_i}{N^i}\right)=\sum\frac{\varphi(m_i)}{N^i},$$

where $m_i=0,1,\ldots,$ or $N-1$: $m_i\not\to 0$; and $N^i=N\cdot N\cdots N$ (i times). Find the subgroup H of the permutation group such that φ^* is continuous if $\varphi\in H$. Further, show that φ^* is differentiable for such φ.

Solution by the St. Olaf Problem Group, St. Olaf College. Suppose there is an integer $i_0\in\{0,1,\ldots,N-2\}$ for which $|\varphi(i_0)-\varphi(i_0+1)|>1$. Then we can construct a sequence $\{x_j\}_{j=1}^{\infty}$ which converges to a number x_0 but whose images $\{\varphi^*(x_j)\}_{j=1}^{\infty}$ do not converge to $\varphi^*(x_0)$. Specifically, let $x_j=\sum_{k=1}^{\infty}n_{jk}/N^k$, where

$$n_{jk}=\begin{cases} i_0+1 & k=1\\ 0 & 1<k\leqq j.\\ 1 & k>j\end{cases}$$

Then $\{x_j\}_{j=1}^{\infty}$ converges to $x_0=\sum_{i=1}^{\infty}m_i/N^i$, where $m_1=i_0$, $m_i=N-1$ for $i>1$. Obviously $\varphi^*(x_j)\not\to\varphi^*(x_0)$ since $|\varphi(n_{j1})-\varphi(m_1)|=|\varphi(i_0+1)-\varphi(i_0)|>1$. Therefore "adjacent" integers in $\{0,1,\ldots,N-1\}$ must be permuted by φ to adjacent integers. But there are only two such permutations: $\varphi(i)=i$ and $\varphi(i)=N-1-i$; these induce, respectively, the mappings $\varphi^*(x)=x$ and $\varphi^*(x)=1-x$, both of which are continuous and differentiable.

Also solved by Allen Beadle, Michael Ecker, Vinod Kumar Grover (India), Wallace Hamilton, R. D. Leitch (England), O. P. Lossers (Netherlands), Nicholas Passell, and the proposer.

Note. Hamilton poses the following as an extension of the problem: Let φ be a binary operation on $\{0,1,2,\ldots,N-1\}$, and let $\varphi^*\sum m_i/N^i=\sum\varphi(m_i,m_{i+1})/N^i$. When is φ^* continuous?

Truncated Exponential-type Series

6056 [1975, 942]. *Proposed by Simeon Reich, University of Chicago*

Let $\{a_n\}$, an increasing sequence of real numbers, tend to infinity and set $p_n(t)=\sum_{k=0}^{n}a_{n-k}t^k/k!$. Is it true that $\lim_{n\to\infty}e^{-a_n}p_n(a_n)/a_n=0$? (Remark: It can be shown, for example, that if $a_n=n$ for all n, then the answer is in the affirmative.)

Solution by Jan Boman, University of Stockholm, Sweden. The statement is false. In fact the limit in question is >0 for any increasing sequence a_n tending to infinity and satisfying (i) $a_n\leqq(1-\delta)n$ for some $\delta>0$, and (ii) $a_{2n}\leqq Ca_n$ for some C. The sequence $a_n=n/2$ satisfies these conditions.

Set $G(x)=\sum_{k\leqq x}x^k/k!$ for real $x\geqq 1$. To prove our statement we need the following (known) estimate

$$(*)\qquad G(x)\geqq ce^x\quad\text{for some}\quad c>0.$$

This estimate can be proved by using Stirling's Formula and comparing the sum with the integral

$$\int_1^x(xe/u)^u u^{-1/2}du>\int_{x/2}^x\ldots du=\sqrt{x}\int_{1/2}^1 e^{xu(1-\log u)}u^{-1/2}du.$$

By a theorem on asymptotic expansions it suffices to note that the function $\varphi(u)=u(1-\log u)$ has the

properties $\max \varphi(u) = \varphi(1) = 1$, $\varphi'(1) = 0$, $\varphi''(1) \neq 0$, in order to conclude that the last integral is of the order, constant $\cdot \exp(x)/\sqrt{x}$ as $x \to \infty$.

Assume that $a_n \leqq (1-\delta)n$. Using first the fact that a_n is increasing and then (*) we get (we write $[x]$ = integral part of x)

$$p_n(a_n) \geqq \sum_{k \leqq (1-\delta)n} a_{n-k} a_n^k / k! \geqq a_{[\delta n]} \sum_{k \leqq (1-\delta)n} a_n^k / k! \geqq c a_{[\delta n]} \exp(a_n).$$

From (ii) it follows that $a_{[\delta n]} \geqq c_1 a_n$ for some $c_1 > 0$. This proves the statement.

Note. Using the estimate $\sum_{k \leqq x/(1+\delta)} x^k/k! = o(e^x)$ as $x \to \infty$ $(\delta > 0)$ we prove that the limit in the problem is in fact zero for every increasing sequence a_n satisfying $a_n \geqq (1+\delta)n$. On the other hand, there exist increasing sequences a_n such that $a_n \geqq n$ and $\limsup_{n\to\infty} \exp(-a_n)p_n(a_n)/a_n > 0$.

I. *The limit in question is zero for every increasing sequence a_n satisfying $a_n \geqq (1+\delta)n$.*

Proof. Since a_n is increasing

$$p_n(a_n) \leqq a_n \sum_{k=0}^{n} a_n^k/k! \leqq a_n \sum_{k \leqq a_n/(1+\delta)} a_n^k/k!.$$

The last sum is $o(\exp(a_n))$ by the estimate stated in the Note. This proves the statement.

II. *There exist increasing sequences a_n such that $a_n \geqq n$ and* $\limsup_{n\to\infty} \exp(-a_n)p_n(a_n)/a_n > 0$.

Proof. We choose a_n constant on large intervals, and the construction is carried out inductively. Assume that $a_1, \ldots, a_m$ are chosen and set $a_n = q$ for $m < n \leqq q$, where q is to be chosen. Then

$$p_q(a_q) = \sum_{k=0}^{q} a_{q-k} a_q^k/k! > \sum_{k=m+1}^{q} \ldots = q \sum_{m+1}^{q} q^k/k! > q \cdot c \cdot e^q$$

for large q and some constant $c > 0$. In the last inequality we have invoked the estimate (*). Since $q = a_q$ this proves the statement.

Also solved by Allen Beadle, Robert Breusch, Paul Bruckman, S. W. Dharmadhikari, Paul Erdös (Israel), William Habakkuk, Daniel Gootkind, Ellen Hertz, A. C. Hindmarsh, Donald Knuth, Joel Levy, O. P. Lossers (Netherlands), William Margulies, L. E. Mattics, A. Meir (Canada), Itrel Monroe, D. J. Newman, Bryce Parry, W. C. Waterhouse, and the proposer.

Editor's comment. Several solvers achieved solutions by using the theory of the Poisson distribution. In particular, if $a_n = n^{1-\varepsilon}$, $1 > \varepsilon > 0$, then the limit is 1.

Determinants of Matrices

6057 [1975, 942]. *Proposed by Anon, Erewhon-upon-Yarkon*

Let A, B, C, D be $n \times n$ matrices such that $CD' = DC'$, where the prime denotes transpose. Prove that

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |AD' - BC'|.$$

I. *Solution by Anand Tamhankar, Buffalo, New York.* It is clear that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} D' & 0 \\ -C' & I \end{pmatrix} = \begin{pmatrix} AD' - BC' & B \\ CD' - DC' & D \end{pmatrix}$$

so that, using $CD' = DC'$ and taking determinants, we have

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} |D'| = |AD' - BC'| |D|.$$

This already gives the result if D is nonsingular. If not, then treating the determinants as functions

of real variables and using continuity gives the result. Otherwise, considering the elements of A, B, C, D as indeterminates generating a polynomial ring over the integers will allow the cancellation of the polynomial $|D| = |D'|$.

II. *Solution by Herbert Carus, Paterson, New Jersey.* The proof follows directly from the theorem for the determinant of a partitioned matrix, viz., if A and D are square and D^{-1} exists, then

$$\Delta = \det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(A - BD^{-1}C)\det D.$$

(See F. R. Gantmacher, *Matrizenrechnung*, 1958, vol. I, p. 43.) In the present case, since A and D are of the same order

$$\begin{aligned}\Delta = \det(A - BD^{-1}C)\det D' &= \det(AD' - BD^{-1}CD') \\ &= \det(AD' - BC'),\end{aligned}$$

by the hypothesis $CD' = DC'$. Since Δ is a continuous function of the elements of D, the requirement that D be invertible may be removed at this point.

The related problem with the hypothesis $CD = DC$ and the conclusion $\Delta = \det(AD - BC)$ occurs in several textbooks, e.g. L. Mirsky, *Linear Algebra*, 1955, p. 110.

Also solved by Mark Balas, Joseph Bastian & Gerd Fricke, D. R. Breach, Paul Bruckman, John Bryant & Robert Gilmer, Francis Callahan, P. G. Chauveheid (Belgium), C. G. Cullen, William Habakkuk & Melvin Hochster, Yasuhiko Ikeda, A. A. Jagers (Netherlands), Shyam Johari, H. Kestelman (England), O. P. Lossers (Netherlands), L. E. Mattics, Robert Spira, John Tung, William Watkins, and the proposer.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN
with the assistance of the mathematics departments of St. Olaf and Carleton Colleges
COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, CARLETON COLLEGE

We invite readers to submit reviews of significant recent college-level mathematics books. We especially encourage reviews based on classroom use, or comparative reviews of several related books. Reviews should ordinarily not exceed two pages (per book) typed double spaced. Manuscripts of reviews as well as books submitted for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057.

Elementary Calculus. By H. Jerome Keisler. Prindle, Weber & Schmidt, Boston, Massachusetts, 1976. xviii + 941 pp. $20.50. (Telegraphic Review, April 1976.)

This review concerns our use of a radically new treatment of 'freshman' calculus which was made possible by Abraham Robinson's recent (1960) discovery of a consistent modern theory of infinitesimal numbers extending the ordinary scalars $\mathbb{R}$. Over the last few years several instructors at Iowa have taught from a preliminary edition of Keisler's book. In 1975 we used the typescript of the recently-published edition.

Keisler has given a simple formulation of Robinson's theory which allows a rigorous but highly intuitive development of calculus à la Leibniz. Keisler's book goes considerably beyond correct interpretations of $(df/dx)(x)$ as nearly a ratio of infinitesimal changes and of $\int_a^b f(x)dx$ as nearly an infinite sum of infinitesimal rectangles — he provides a technique which allows the instructor to easily give complete, rigorous, intuitive proofs, even of such results as the intermediate and extreme value

theorems (which tend to have particularly awkward epsilon-delta proofs). Although we feel that freshman calculus should not emphasize proofs, with this method one does them quickly but correctly and moves on. More importantly, Keisler's method permits precise, intuitive geometric derivations of such results as the derivatives of sine and cosine by nearly similar triangles — one finite, one infinitesimal. (He also includes the addition formula approach.) Infinitesimals provide a method for *calculating* limits as well. (Epsilon and delta methods require the answer in advance.)

We believe that the main importance of calculus is what the name implies — *calculation.* Calculation of areas, lengths, volumes, tangents, maxima, rates of change, sums of series, and so forth, especially by use of the derivative and integral. In this regard, Leibniz' approach to calculus has long held a recognized advantage at least in guiding correct calculation; his notation persists even in the most pedantic epsilon-delta courses, as it should. The infinitesimal approach permits the most direct interpretation of the calculations. A student who refuses to draw little disks or shells when computing volumes of revolution ultimately encounters difficulty and usually dislikes the problems. Keisler's approach makes the 'sum of infinitesimal disks' heuristic nearly true (within an infinitesimal). One can think of the small disks as part of a limiting process, but the infinite sum of infinitesimals Leibniz had in mind (in the form provided by Robinson–Keisler) is just a very good approximation very far along the process, it is more direct but not at odds with finite limits.

Moreover, Keisler does a very careful job of showing the student the epsilon-delta or limit formulations of the concepts of infinitesimal calculus when he discusses numerical approximation, for example, the trapezoidal rule. We believe that a student who has gone carefully thru Keisler's chapter 5 will know as much about "epsilon-delta" as the students we have taught using Apostol's *Calculus* (2nd edition) while the Keisler student should have a stronger intuitive feeling for the integral and derivative because of the infinitesimals. On the other hand, error estimates are not easy. Even in our computer lab offered in conjunction with calculus they meet with limited success. Therefore, good heuristic arguments which infinitesimals can now provide rigorously are important in any approach. We think it is desirable to have intuition and rigor more closely linked than is possible in conventional approaches at the usual freshman level.

Apostol's *Calculus* is a very good book and to our mind this is largely because of a feature it has in common with Thomas's *Calculus and Analytic Geometry* (we have used the 4th edition) — abundant well-chosen problems of various difficulties beginning with simple exercises. The old preliminary edition of Keisler's book which we used for several years was inadequate in this regard, but the final version is comparable to Thomas's, has accurate answers, good worked examples, and only slightly fewer difficult theoretical problems than Apostol's. Chapter one of Keisler's book has one important problem set which both of the others lack — word problems which ask the students for the appropriate formulas (from geometry and so on). Keisler slips this in before calculus under the guise of studying extension of formulas to the hyperreals, but he certainly has in mind pedagogically separating the difficulty students encounter translating English into mathematicalese from calculus per se. It is very helpful. The basic exercises of Apostol, Keisler, or Thomas represent the core of the calculus. Euler could have solved them without the rigorous interpretations of either Weierstrass or Robinson. Keisler's students can understand them in terms of intuitive infinitesimals. Better students will even overcome some of the mysteries they presented for Euler or Cauchy when, for instance, they equated the notions "is infinitesimally close to" and "is equal to" or added infinitely many infinitesimals and thought the answer should be infinitesimal. This core understanding of calculus is mostly a function of how many problems the students ponder and solve for themselves — Keisler provides the medium and a good deal more.

One of the important pedagogical innovations in Keisler's book is his use of heuristic "infinitesimal microscopes" and "infinite telescopes" (these can be precisely defined by infinitesimal similarity transformations and infinite translations with range in the unit disk viewed modulo infinitesimals). These provide a simple way to explain many of the difficulties classically associated with the formulas in calculus, so they are a strong point in his approach. A simple example of their use is to help derive

the arc length formula for a curve. In a common mistake one only cuts the curve $y = f(x)$ along the x-axis, but under an infinitesimal microscope (Figure 1) we see:

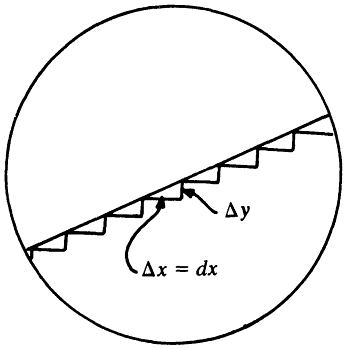


FIG. 1

and certainly we cannot ignore the Δy-terms. This example also brings up two other especially elegantly formulated basic lemmas in Keisler's book, the increment theorem and the infinite sum theorem. The increment theorem says 'under an infinitesimal microscope a smooth curve is indistinguishable from a straight line'. In other words, we cannot see the difference between Δf and df even if we magnify by $1/dx$ for dx infinitesimal and positive. As a formula, $f(x+dx)-f(x)= f'(x)dx + \varepsilon \cdot dx$, where ε is infinitesimal. With two infinitesimal microscopes, focusing one off center in the eyepiece of the other, we can see ε as in Figure 2. (We hasten to add that at first Keisler simplifies the increment theorem more than we have done here.)

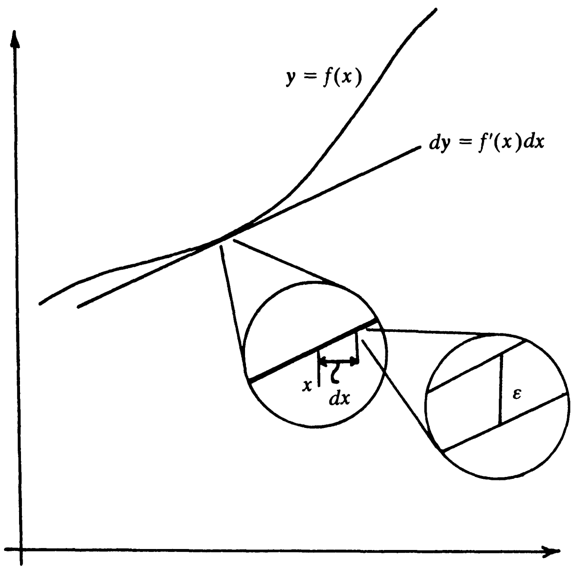


FIG. 2

Now this means that the Pythagorean theorem nearly applies (in the first microscope) to the Δx and Δy of the arc length above to yield "*Δl is infinitesimally close to*

$$\sqrt{(dx)^2+(dy)^2},$$

compared with dx." The infinite sum theorem then assures us that

$$l = \int_a^b \sqrt{(dx)^2 + (dy)^2},$$

that is, the infinite sum of infinitesimal errors between Δl and $\sqrt{(dx)^2 + (dy)^2}$ only adds up to an infinitesimal. (In general, $1/H + 1/H + \cdots + 1/H$, H-times, is one, by the solution axiom, even when H is infinite and $1/H$ infinitesimal. Even worse, $1/H + 1/H + \cdots + 1/H$, H^2-times, is H, an infinite number.) The shell method provides another nice application of the infinite sum theorem, because a $(dx)^2$-term sums to an infinitesimal and therefore does not appear in the integral. The first calculus text, *Analyse des infiniment petits pour l'intelligence des lignes courbes* (1696), by de l'Hospital, contains as one of its most fundamental axioms that curves were made of infinitesimal polygons. The increment theorem is a modern formulation of that axiom.

We highly recommend Keisler's book to interested instructors for "Engineering Calculus" through "Honors Calculus," with the change in emphasis coming from the instructor. Keisler's book gives correct intuitive proofs, once the instructor has helped the student over the axiom hurdle in chapter one, that is, once the student understands computationally and intuitively what infinite and infinitesimal numbers are. More or less emphasis on theoretical details and choice of topics make the book suitable for the full spectrum of calculus courses. Since his proofs are intuitive, a sketch provides the background for theorems in an intuitive course. Keisler's book has a large number of population and economics problems as well, so it may be interesting for "Liberal Arts Calculus."

We qualify our recommendation of Keisler's book. There are pitfalls, most notably the axioms in chapter one (comparable to Apostol's axiomatics). The book is revolutionary and we feel that instructors should learn the meaning of the axioms well before they attempt to teach them. The axioms need to be played down, explained intuitively in terms of pictures and examples, and dispensed with within a few days (unless students really need the excellent review of algebra also contained in chapter one). Not all the proofs of the basic properties need to be given in class. There is a great temptation for the instructor to apologize for, explain in excess, and philosophize about something which we find the students readily accept — *infinitesimals*. (They don't know "the reals are categorical.")

At the level of chapter one, the axioms say that algebra with infinite and infinitesimal numbers is just like ordinary algebra (in particular, $2 \cdot \delta \neq \delta$ so there are lots of infinitesimals) and some intuitively plausible rules about finite, infinite, and infinitesimal apply (for example, $1/\delta$ is infinite when δ is a nonzero infinitesimal or $\delta \cdot a$ is infinitesimal when a is finite and δ infinitesimal). There are no general laws to evaluate 'infinitesimal times infinite', so the calculations involving ratios of infinite and infinitesimal quantities, products of infinite times infinitesimal, and so on, are an important part of chapter one. The instructor should go over these calculations carefully, but otherwise should not unnecessarily delay introduction of concepts of calculus by laboring axiomatics.

Keisler's chapter two is helpful in this regard in that it postpones continuity in favor of derivatives. The function extension and solution axioms can be returned to as deeper applications are needed. Some occasions of this are simple and natural, for example, extending the sine and cosine to hyperreals and showing they represent projections from the unit hypercircle. Others are more difficult and less direct, for example, explaining precisely what infinite hyperintegers are in order to use them in proofs of the intermediate value theorem or infinitesimal Riemann sums. We do not think it is necessary to go into great detail on these occasions — essentially the axioms say you *can* do what you would like to intuitively. However, an instructor who does not already know about infinitesimals may very well want to examine such questions in detail personally, well in advance of giving intuitive sketches in class.

We do not recommend forcing this book on anyone who is reluctant to learn about infinitesimals. (We don't recommend forcing Apostol or Thomas on instructors either.) Keisler has prepared a short

instructor's supplement, *Foundations of Infinitesimal Calculus*, Prindle, Weber & Schmidt (paper), 1976.

The direct approach to the natural exponential function in chapter 8 is what you always wanted to tell your students, but simply couldn't manage technically using epsilons and deltas — hyperrational numbers will let you do it, but it still isn't easy! Perhaps an instructor might want to use an integral approach to the natural log and spend more time on elementary differential equations or some other topic. Either way we highly recommend the quick introduction to transcendental functions in chapter 2. The book may seem somewhat weak on elementary differential equations. An instructor might want to hand out an extra problem set on second order constant coefficient equations. Elementary complex variables and Euler's equation might be added at the same time. Some of us would also like to see more differential geometry (of curves at least). However, there is a very full year's work of excellently treated topics in the first nine or ten chapters. The remainder would fill a third semester (although we have not yet used the several variables part).

We like the book and will use it again. Our students have reacted favorably to the book. It is written for students to read, including warnings of common errors and outlines of general procedures for students to apply. K. Sullivan has made a study of use of the preliminary edition at a number of schools (this MONTHLY, 83 (no. 5), 1976, pp. 370–375). We believe the book more than fulfills the author's claim — his calculus with infinitesimals teaches the students everything the traditional students learn and it does so very well. It also teaches one more thing — what infinitesimals are. We believe that the latter further aids understanding the calculus and is interesting in its own right.

E. W. MADISON and K. D. STROYAN, University of Iowa

The Foundations of Geometry and the Non-Euclidean Plane. By George E. Martin. IEP, Dun-Donnelly, New York, 1975. xvi + 509 pp. $15.00. (Telegraphic Review, October 1976.)

This is not only the best introduction to the topics in the title, it is one of the best advanced undergraduate textbooks I have read or used. The author is original, accurate, tremendously informative, witty, and readable. After a four-chapter introduction (logic, sets, the reals, incidence structures), Birkhoff's axiomatization of absolute geometry (straightedge, ruler, scissors, protractor, mirror) is presented, with many side trips such as taxicab geometry, the elements, Hilbert's and Pieri's systems. Absolute geometry is then pursued for over one hundred pages, including reflections, circles, Saccheri's theorems, and biangles. In the non-euclidian geometry course at Hofstra we ran barefoot through this vast field, stopping from time to time to admire a particularly lovely flower, then slowed to a sedate march (appropriately shod) beginning with the author's "Let's do it!" (p. 334) after which the hyperbolic parallel postulate, his sixth and last axiom, is laid down. There follows a one hundred sixty page development of hyperbolic geometry including the classification of isometries, a complete treatment of trigonometry and the fundamental formula ($\cos \pi(x) = \tanh x$), and ending with a proof that with hyperbolic instruments in the hyperbolic plane the circle *can* be "squared," although, of course, there are no squares in the hyperbolic plane.

The author is not afraid to give details. Leaving messy proofs as exercises, misuse of the words "similarly" and "trivial" and other types of drivel sometimes found in textbooks do not occur here. The few errors noted are relatively harmless. There are more than 650 exercises, ranging from very easy to at least fairly hard. All but four of the thirty-four chapters end with a collection of "Graffiti," which skip around among sagacious quotes, logical puzzles, humorous quotes, interesting formulae, and nice frieze patterns. While the author may be said to rival the great H.S.M. Coxeter at this sort of thing, he cannot be said to surpass him.

If you get a chance to teach foundations of geometry, or non-euclidean geometry to advanced undergraduates, be sure to consider this book. Otherwise, try browsing in it just for fun.

ROBERT J. BUMCROT, Hofstra University

TELEGRAPHIC REVIEWS

Telegraphic reviews are designed to give prompt notice of new books with sufficient information to assist our readers in deciding whether to order an examination copy or to suggest library purchase. Possible uses are indicated as follows:

T = textbook P = professional reading
S = supplementary reading L = undergraduate library purchase
13 to 18 = freshman to second year graduate level usage
1 to 4 = appropriate time in semesters to cover text

Asterisks (*) or question marks (?) denote special positive or negative emphasis, respectively. Publishers are denoted by standard abbreviations; complete addresses may be found in *Books in Print*.

GENERAL, P**, L*, *A Basic Library List for Four-Year Colleges, Second Edition*. MAA, 1977, 106 pp, $4.50 (P). 700 titles, including recommendations for a 300 book nucleus library, selected from the 1965 *First Edition* and the 7000-plus college-level mathematics books published since then. This edition is twice as large as the original and emphasizes recently developing areas of the mathematics literature (combinatorics, computing, applications). Every college *librarian* should have this *List* and every college library should have its contents or their equivalent. LAS

GENERAL, S(13-16), L*, *Ten Faces of the Universe*. Fred Hoyle. Freeman, 1977, ix + 207 pp, $10.95; $6.95 (P). A vivid, personal portrait, ranging from primeval elements of matter and force, pattern and probability to pessimistic speculation on the impending break in the population wave due to interaction between industrial productivity and human procreation. Hoyle captures the universe in an imaginative tapestry, woven from the space-time threads of its constituent particles, and organized by mathematical law. LAS

GENERAL, S*, L*, *Chess Skill in Man and Machine*. Ed: Peter W. Frey. Springer-Verlag, 1977, xi + 217 pp, $14.80. Programming a computer to play good chess has become a central problem in artificial intelligence--of interest to psychologists, computer scientists, and chess players. Taken together, these eight informative essays nicely summarize the current state of the art after 20 years of research. The major discovery is that the problem is much more difficult than originally believed--inexorably related to basic problems regarding human thinking. The flavor is clear: the challenge is here and the journey has only begun. LCL

GENERAL, S*(11-15), L**, *1001 Problems in High School Mathematics, Book 1, Problems 1-100, Solutions 1-50 (Preliminary Edition)*. Ed: E. Barbeau, M. Klamkin, W. Moser. Canadian Math Congress (3421 Drummond St., Montreal, Canada H3G 1X7), 1977, 56 pp, $1.80 (P). The first in a planned series of booklets to be issued every six months, containing a selection of interesting problems (arranged in no particular order) that can be solved by ingenious elementary methods. Includes an appendix giving a Tool Chest of handy devices for cracking these tough nuts. LAS

GENERAL, S*, L***, *Did You Say Mathematics?* Ya. Khurgin. Trans: George Yankovsky. MIR (US Distr: Imported Pub, 320 W. Ohio St., Chicago, IL 60610), 1974, 360 pp, $1.75 (P). A compact, informal, popular introduction to selected topics (e.g., rubber-sheet mathematics, the mathematics of a saddle, codes, probability theory) written by a versatile applied Soviet mathematician (radiophysics, cybernetics, neurophysiology, psychiatry) "for those who are separated from mathematics by a wall of formulas, equations, proofs and graphs." A Russian equivalent of Courant and Robbins' famous *What is Mathematics?*, the monograph appears prosaic and proletarian, not likely to excite the Sesame Street generation. Yet for its price it is a good buy. LAS

BASIC, T?, *Principles of Counting*. Amy Pohl. Wills Pub, 1977, viii + 72 pp, $1.95 (P). Permutations and combinations culminating with the binomial theorem (no proof). Excessive detail, dull problem sets, misdirected emphases. LCL

PRECALCULUS, T(13), *Trigonometry for College Students*. Karl J. Smith. Brooks/Cole, 1977, xi + 283 pp, $11.95. Rather traditional treatment of plane trigonometry. Somewhat carelessly written, and not always mathematically sound. JD-B

PRECALCULUS, T(13: 1), S*, L*, *The Functions of Algebra and Trigonometry*. Kenneth P. Bogart. HM, 1977, xvi + 512 pp, $13.50. Discusses properties of functions in general (e.g., composition, inverses) with a full treatment of all the elementary functions. Includes finite mathematics. Best suited for a course between precalculus algebra and calculus (or good supplementary reading for calculus). Exposition is very good. Consistent effort to give applied examples and problems. CB

PRECALCULUS, S(13), *Concise Review of Algebra and Trigonometry*. A.W. Goodman. Saunders, 1977, 139 pp, $4.95 (P). A brief review of essentials which can accompany any calculus book. Should be a large audience for this. LLK

PRECALCULUS, T(13: 1), *College Algebra*. J.S. Ratti. Macmillan, 1977, vi + 282 pp, $12.95. Succinct presentation--algebra, equations, functions (exponential, logarithmic, but no trigonometric), matrices, complex numbers, induction, sequences and progressions, binomial theorem, permutations and combinations. An encouraging collection of word problems. Attractive format. LCL

EDUCATION, T*(1, 2), S, L, *Contemporary Mathematics, Second Edition*. Bruce E. Meserve, Max A. Sobel. P-H, 1977, xiii + 592 pp, $12.95. Major changes from first edition (ER, January 1975) include: new chapters on metric measurement; computers and calculators; expansion of "Pedagogical Explorations"; addition of "Readings and Projects", a useful guide to relevant literature, especially that found in *The Arithmetic Teacher*, *The Mathematics Teacher* and the NCTM yearbooks. LCL

EDUCATION, S(7-12). *Adventures With Your Hand Calculator.* Lennart Råde, Burt A. Kaufman. CEMREL, 1977, iv + 131 pp, $4.95 (P). 20 areas of numerical exploration (primes, random digits, magic squares, etc.) requiring only rudimentary calculators. Extensive commentary in the second half provides insight and hints (with occasional references) for more extensive projects. LAS

EDUCATION, P*, L. *Math Activities for Child Development, Second Edition.* Enoch Dumas, C.W. Schminke. Allyn, 1977, xiv + 341 pp, $6.95 (P). A compendium of nearly 500 activities, games, projects, puzzles and resource materials for mathematics enrichment in the elementary school. LAS

HISTORY, S*, P, L**. *Mathematics and Mathematicians.* P. Dedron, J. Itard. Trans: J.V. Field. Transworld Pub, 1973. *V. 1,* 325 pp, £1.50 (P); *V. 2,* 222 pp, £1.00 (P). A translation of the 1959 *Mathématiques et Mathématiciens* prepared especially for the Open University's course in History of Mathematics. It presents important pre-nineteenth century mathematics in a scholarly yet mathematically elementary manner, quoting extensively from primary sources and relating mathematical to historical contexts. V. 1, a general outline of the growth of mathematics in western Europe, traces "intentions, hesitations and doubts" which led to the flowering of new methods during and after the Renaissance. The second volume treats in detail several particular problems (calculation, first and second degree problems, trisection, etc.) as case studies in method. LAS

HISTORY, P, L*. *Selected Papers of Alfréd Rényi, V. I-III.* Ed: Pal Turán. Akadémiai Kiadó, 1976. *V. I,* 628 pp; *V. II,* 646 pp; *V. III,* 667 pp, $120. Approximately half of Rényi's papers (with Hungarian, Russian and Chinese works translated into English) arranged chronologically (1948-1956; 1956-1961; 1962-1970) with frequent notes and cross references to related literature. Each volume begins with a complete Rényi bibliography (of 355 papers) and a brief biography. LAS

HISTORY, P, L**. *Greek Mathematical Thought and the Origin of Algebra.* Jacob Klein. Trans: Eva Brann. MIT Pr, 1968, xv + 360 pp, $4.95 (P). A thorough scholarly study, emphasizing classical Greek roots (especially in Plato), Diophantus, and François Viète; includes a translation of Viète's *Introduction to the Analytical Art.* A paperback edition of the 1968 hardcover English translation (TR, June 1969; ER, May 1971) of the 1934-36 German original. LAS

COMBINATORICS, P, L. *The Theory of Partitions.* George E. Andrews. A-W, 1976, xiv + 255 pp, $16.50. A thorough look at the number theoretic and combinatorial aspects of partition theory. Among the important topics covered are infinite series and infinite product generating functions, Rogers-Ramanujan identities, sieve methods, higher dimensional partitions, partitions and finite vector spaces and sets. Contains a number of interesting exercises, historical notes, and extensive bibliography. An excellent text and a valuable reference. SG

NUMBER THEORY, T(15: 1, 2), S*, P, L*. *Elementary Number Theory, A Problem Oriented Approach.* Joe Roberts. MIT Pr, 1977, vi + 375 pp, $12.50 (P). Much of this book consists of sets of problems which lead to well known theorems. Several topics are included which do not generally appear in beginning number theory texts. A complete set of solutions and a substantial list of references are included. This hand-calligraphed book is a thing of beauty. CEC

LINEAR ALGEBRA, S(13). *An Introduction to Applied Linear Algebra.* Norman Locksley. Wills Pub, 1977, v + 106 pp, $3.50 (P). A concise treatment of matrices and linear programming. Neat, compact, and honest. LLK

LINEAR ALGEBRA, T(14: 1). *Linear Algebra, Second Edition.* Michael O'Nan. HarBrace J, 1976, xi + 335 pp, $11.95. Changes from the first edition (TR, May 1971; ER, June 1973) include more examples and some new exercises, extension of chapter on eigenvalues, and more extensive index. Still very short on answers to selected exercises in back of book. LLK

ALGEBRA, S(18), P. *Near-Rings, The Theory and its Applications.* Günther Pilz. Math. Stud., V. 23. North-Holland, 1977, xiv + 393 pp, $24 (P). A monograph on near-rings, i.e., rings except that addition is not assumed commutative and only one-sided distributivity is postulated. Exhaustive bibliography. JD-B

ALGEBRA, P. *Lecture Notes in Mathematics-545: Noncommutative Ring Theory.* Ed: J.H. Cozzens, F.L. Sandomierski. Springer-Verlag, 1976, iv + 212 pp, $10.20 (P). Eight papers, including lots of open problems. LCL

ALGEBRA, T(16-18), S, L. *The Theory of Groups, Second Edition.* Marshall Hall, Jr. Chelsea, 1976, xiii + 434 pp, $9.95. Unaltered reprint of the 10th (1968) printing of the well-known original 1959 Macmillan edition. A ten chapter course (with exercises)followed by ten chapters on special topics. LAS

ALGEBRA, P. *Lattice Theoretic and Logical Aspects of Elementary Topoi.* Christian Juul Mikkelsen. Aarhus U, 1976, iv + 122 pp, (P). An elementary topos is a category which looks like the category of sets. This monograph applies techniques of lattice theory to the study of topoi. PJM

ALGEBRA, P. *Transformation Groups.* Ed: Czes Kosniowski. London Math. Soc. Lect. Notes., No. 26. Cambridge U Pr, 1977, vii + 306 pp, $8.95 (P). Proceedings of an August 1976 conference at the University of Newcastle-upon-Tyne, including ten research papers plus 15 short summaries. LAS

ALGEBRA, P, L. *Lecture Notes in Mathematics-554: Mal'cev Varieties.* Jonathan D.H. Smith. Springer-Verlag, 1976, viii + 158 pp, $7.40 (P). An algebra is a set X with a collection of operations $X^n \to X$. A set of algebras with the same operations is a variety. A Mal'cev variety is a nice variety: examples: groups, rings, modules over a ring, Lie algebras. This monograph studies the properties common to all Mal'cev varieties. PJM

ALGEBRA, T(16-18: 1), S, L. *Lectures on Rings and Modules, Second Edition.* Joachim Lambek. Chelsea, 1976, viii + 183 pp, $7.95. Corrected reprinting of the 1966 first edition (TR, April 1969; ER, November 1969). Treats associative rings (with units) and their modules. LAS

ALGEBRA, S(17-18), P. *Orderable Groups.* Roberta Botto Mura, Akbar Rhemtulla. Lect. Notes in Pure and Appl. Math., V. 27. Dekker, 1977, iv + 169 pp, $19.75 (P). On certain modern developments in the theory of orderable groups. Concentrates on relations between order properties and such group-theoretic conditions as nilpotency, solvability and finiteness of rank. JD-B

ALGEBRA, S(17-18), P. *Ring Theory II, Proceedings of the Second Oklahoma Conference.* Ed: Bernard R. McDonald, Robert A. Morris. Lect. Notes in Pure and Appl. Math., V. 26. Dekker, 1977, xviii + 295 pp, $25 (P). Fourteen expository papers on recent developments in the theory of commutative and noncommutative rings. JD-B

CALCULUS, T(13-14: 2). *Calculus: An Applied Approach.* Thomas Wonnacott. Wiley, 1977, xiv + 514 pp, $14.50. An informal and consistently unrigorous introduction to the calculus of one and several variables. A good deal on finite differences and a chapter on linear differential equations, but almost nothing on series and no mention of Mean-Value Theorem. Examples chiefly from business or economics. JD-B

CALCULUS, S(13). *Problems in Calculus and Analytic Geometry.* Richard J. Palmaccio. J. Weston Walch, Pub, 1977, iv + 148 pp, $3.50 (P). About 60 problems and detailed solutions, mostly applied, requiring "integration" of traditional calculus topics. A useful supplement for teachers, although good "applied" calculus books contain hundreds of similar problems. LAS

CALCULUS, T(15-16: 1, 2), P, L. *Functions of Several Variables, Second Edition.* Wendell Fleming. Springer-Verlag, 1977, xi + 411 pp, $16.80. Changes from the 1965 *First Edition* include a new chapter on elementary topology, additional applications to thermodynamics and mechanics, and new proofs of the inverse function and divergence theorems. LAS

REAL ANALYSIS, S(17), P. *Subharmonic Functions, V. I.* W.K. Hayman, P.B. Kennedy. Acad Pr, 1976, xvii + 284 pp, $25.50. London Mathematical Society Monograph on subharmonic functions in R^m, $m \geq 2$. Development begins from the definition. From the preface: "...steers an intermediate course between books on various aspects of abstract potential theory and books on function theory...". CB

COMPLEX ANALYSIS, P. *Algebraic Methods in the Global Theory of Complex Spaces.* Constantin Bănică, Octavian Stănăşilă. Wiley & Edit. Acad. (Romania), 1976, 296 pp, Lei. 29. "The book is intended mainly for experts in complex spaces who are not fully acquainted with the algebraic aspect, and for experts in algebraic geometry who wish to be introduced to the theory of complex spaces." CB

NUMERICAL ANALYSIS, S(17-18), P. *Approximation auf dem kubischen Gitter.* S.G. Michlin. Math. Reihe, B. 59. Birkhäuser, 1976, ix + 194 pp, sFr. 38. A monograph for specialists, devoted largely to the work of the author, on the finite-element method. JD-B

FUNCTIONAL ANALYSIS, T(16-18: 1), S, L*. *Introduction to Hilbert Space, Second Edition.* Sterling K. Berberian. Chelsea, 1976, xi + 206 pp, $7.50. Essentially unaltered reprint of the 1961 Oxford U. Pr. edition. LAS

OPTIMIZATION, P, L. *Lecture Notes in Economics and Mathematical Systems-128: Integer Programming and Related Areas, A Classified Bibliography.* Ed: C. Kastning. Springer-Verlag, 1976, xii + 495 pp, $15.20 (P). A computer-coded index of [illegible] publications (through 1975) related to integer programming. Full entries are listed alphabetically by first author; then abbreviated entries are listed under 41 subject headings; finally an author index keys coauthors to the primary alphabetical list. The editor intends to maintain the computer base with current articles. LAS

ANALYSIS, S(17-18), P. *Lecture Notes in Mathematics-562: Nilpotent Lie Groups: Structure and Applications to Analysis.* Roe W. Goodman. Springer-Verlag, 1976, x + 210 pp, $10.20 (P). Deals with recent results, due largely to Stein and his collaborators, on uses of nilpotent Lie groups in the representation theory of semi-simple Lie groups, complex analysis, and partial differential equations. JD-B

ANALYSIS, T**(16-17: 1), S**, P, L***. *Applied Nonstandard Analysis.* Martin Davis. Wiley, 1977, xii + 181 pp, $16.95. A beautiful development placing key stress on the Transfer Principle. Written by a leading mathematics expositor, this introduction leads the beginner through the necessary logic, and, assuming only minimal background in algebra and analysis, procedes to apply the "nonstandard" method to real analysis, topology, and Hilbert space. Nothing new here for the specialist. LCL

ANALYSIS, S*(14-16). *Problems and Theorems in Analysis, V. I: Series, Integral Calculus, Theory of Functions.* G. Pólya, G. Szegö. Trans: D. Aeppli. Springer-Verlag, 1972, xix + 389 pp, $9.80 (P). Paperback "study edition" of the 1972 English edition (TR, January 1973). LAS

GEOMETRY, S**(13-16), P**, L**. *Geometry, Relativity and the Fourth Dimension.* Rudolf v.B. Rucker. Dover, 1977, 133 pp, $2.75 (P). An extraordinary exploration of the geometry of space-time, beginning with elementary concepts of four-dimensional geometry, and concluding with speculation (based on the most current work in theoretical physics) concerning the geometric nature of fundamental physical reality. Not just another exposition of special relativity, but an intriguing, partly original synthesis of geometry and physics, of fact and fantasy. LAS

TOPOLOGY, P. *Lecture Notes in Mathematics-557: Smooth S^1 Manifolds.* Wolf Iberkleid, Ted Petrie. Springer-Verlag, 1976, 163 pp, $7.40 (P). A study of manifolds with smooth actions $S^1 \times M \to M$. Part I is algebraic tools, Part II applies them. PJM

TOPOLOGY, P. *The Quantitative Theory of Foliations.* H. Blaine Lawson, Jr. CBMS Reg. Conf. in Math., No. 27, AMS, 1977, v + 65 pp, $6.40 (P). A foliation is a decomposition of a manifold into lower dimensional manifolds. Given a manifold, one can ask how many distinct foliations does it have. Results towards an answer are presented in this booklet. PJM

TOPOLOGY, T(16-17: 2), S, L. *Lecture Notes on Elementary Topology and Geometry.* I.M. Singer, J.A. Thorpe. Springer-Verlag, 1967, viii + 232 pp, $14.80. Reprint (in Springer's new *Undergraduate Texts* series) of the 1967 Scott Foresman original paperback edition (TR, November 1967; ER, February 1968). LAS

TOPOLOGY, T*(16: 1, 2), S*, L**. *Graphs, Surfaces and Homology.* P.J. Giblin. Halsted Pr, 1977, xv + 329 pp, $10.50 (P). An excellent, reasonably priced introduction to algebraic topology really suited to undergraduates. After chapters on graphs and closed surfaces, the author introduces the machinery of homology theory and applies it backwards to some of the earlier results and forwards to a study of imbedding graphs in surfaces. No exercises as such, but examples are worked out in different amounts of detail--the relatively unworked ones are nice exercises. No point set topology required as a prerequisite. Necessary abelian group theory is contained in an appendix. PJM

PROBABILITY, P*. *Inequalities for Stochastic Processes (How to Gamble if You Must).* Lester E. Dubins, Leonard J. Savage. Dover, 1976, xiv + 255 pp, $4 (P). Unabridged paperback republication, with corrections and a new Bibliographic Supplement and Preface, of the authors' well-known 1965 McGraw-Hill book. (The title and subtitle have been permuted "to alleviate the present publisher's concern about possible misunderstandings as to the nature of the book.") A technical treatise on gambling strategies, viewed as finitely additive, time-discrete stochastic processes. RSK

PROBABILITY, T(13: 1). *The Computation of Probability with Basic, Pilot Edition.* Basil E. Gala. Holden-Day, 1977, iii + 179 pp, $8.95 (P). Integrated introduction to elementary discrete probability and computer programming. Coverage of each is minimal, explanations are sparse, and there are the usual typographical errors. However, the material is nicely integrated. Most sections conclude with a "programmed study set of questions", with answers, and a set of exercises. RSK

PROBABILITY, T(13). *Probability: A Set Theory Approach.* Amy Pohl. Wills Pub, 1977, viii + 218 pp, $3.95 (P). Probabilities by listing possible outcomes, counting techniques, conditional probability, binomial distribution (some confusion about random variables in the definition of expected value). LCL

STATISTICS, T(17: 2). *Mathematical Statistics, Basic Ideas and Selected Topics.* Peter J. Bickel, Kjell A. Doksum. Holden-Day, 1977, xv + 493 pp, $18.95. Presumes a course in probability theory and a "good mathematics background" (linear algebra and advanced calculus, but no measure theory). Theoretical treatment with many exercises and problems. In addition to a detailed treatment of estimation and, to a lesser extent, hypothesis testing, topics covered include general linear models, analysis of discrete data, nonparametric models and decision theory. RSK

STATISTICS, P*. *Prediction Analysis of Cross Classifications.* David K. Hildebrand, James D. Laing, Howard Rosenthal. Wiley, 1977, xv + 311 pp, $22.50. In the Wiley Series in Probability and Mathematical Statistics. Concerned with assessing how well states of one variable can be predicted from one or more other variables, by using a modification of elementary logic obtained by introducing a connective to indicate "tends to be sufficient for", it presents a more refined measure than traditional chi-square for analyzing categorical data. Clearly written with many illustrative examples and a good bibliography. RSK

STATISTICS, P. *Selected Tables in Mathematical Statistics, V. IV, Dirichlet Distribution-Type 1.* Milton Sobel, V.R.R. Uppuluri, K. Frankowski. AMS, 1977, x + 309 pp, $18. Fourth in a series of specialized statistical tables sponsored by the Institute of Mathematical Statistics. Contains tabulated values in a variety of forms of the incomplete Type 1-Dirichlet integral, which is a direct generalization of the incomplete beta distribution for the multinomial case. Introductory material includes ways to use the tables to solve numerous multinomial problems. RSK

COMPUTER SCIENCE, P, L. *Advances in Computers, V. 15.* Ed: Morris Rubinoff, Marshall C. Yovits. Acad Pr, 1976, xiii + 301 pp, $28.50. Five essays describing current issues in automatic programming, algorithm selection, parallel processing, language acquisition and computer-based education. LAS

COMPUTER SCIENCE, S(15-16), P, L. *Simulation with GPSS and GPSS V.* P.A. Bobillier, B.C. Kahan, A.R. Probst. P-H, 1976, xvi + 495 pp, $18.50. Introductory sections on simulation and model building. A gentle development of GPSS: basic constructs, the complete language, special features. Numerous case studies; e.g., warehousing, transportation, job shop, teleprocessing. Not intended as a manual. Directed toward appliers. RWN

COMPUTER SCIENCE, P. *Large Scale Computer Architecture: Parallel and Associative Processors.* Kenneth J. Thurber. Hayden, 1976, 324 pp, $18.95. Concepts of parallel and associative processors for single instruction stream multiple data stream architectures. Case studies of Illiac IV, Pepe, Staran, and Omen. Applications of performance evaluation techniques. Exercises. Bibliography. RWN

COMPUTER SCIENCE, P. *Integrity and Recovery in Computer Systems.* Terry Gibbons. Hayden, 1976, 137 pp, $9.95. Designing software systems with an attempt to achieve operational reliability. Discusses the sources of problems and strategies and techniques for prevention and recovery. RWN

COMPUTER SCIENCE, T(15: 1), S, P, L. *Revised Report on the Algorithmic Language Algol 68.* Ed: A. van Wijngaarden, *et al.* Springer-Verlag, 1976, 236 pp, $9.90 (P). The definitive work on Algol 68 which is a milestone in the development of programming languages. Includes basic definitions, fundamental constructs, context-dependent rules, elaboration-independent constructions, environmental considerations and examples. (1st edition TR, January 1970.) RWN

COMPUTER SCIENCE, P. *Lecture Notes in Computer Science-37: λ-Calculus and Computer Science Theory*. Ed: C. Böhm. Springer-Verlag, 1975, xii + 370 pp, $13.20 (P). Proceedings of a March 1975 symposium in Rome, bracketted by papers of Dana Scott providing context for open problems common to the λ-calculus and theoretical computer science. LAS

APPLICATIONS, T*(15-17: 1), S, L. *Mathematical Modelling*. Ed: J.G. Andrews, R.R. McLone. Butterworths, 1976, xvii + 260 pp, $7.95 (P). Seventeen independent (and independently authored) chapters introduce a variety of real-world mathematical models, for example, steering problems, molecular structure, traffic situations, network flow. Each chapter contains closed and open-ended problems, as well as references for further details. Prerequisites are listed for each chapter, and generally range no higher than sophomore-junior courses. An appealing book for a seminar on mathematical modelling. LAS

APPLICATIONS, T(14), S. *Chemins et flots, ordonnancements*. Robert Faure, Catherine Roucairol, Pierre Tolla. Gauthier-Villars (US Distr: SMPF, 111 W. 57th St., N.Y. 10017), 1976, 233 pp, 49F (P). Applications of graph theory in operations research: travelling salesman, flows in networks, Pert and Cpm ("Ordonnancements"). No exercises, but lots of examples. The French is straightforward. PJM

APPLICATIIONS (ENGINEERING), P, L. *Handbook of Circuit Analysis, Languages and Techniques*. Ed: Randall W. Jensen, Lawrence P. McNamee. P-H, 1976, xxii + 809 pp, $34.50. Includes reference material on nine circuit analysis programs: Astap, Belac, Circ, Circus2, Ecapll, Lisa, Martha, Sceptre, Syscap. Presents the various program capabilities, their practical applications, and some modeling techniques. Chapter references. Appendices. RJA

APPLICATIONS (OPERATIONS RESEARCH), S(15-16). *Economics and Operational Research*. M.H. Beilby. Acad Pr, 1976, x + 174 pp, $12. Written at an introductory level. Emphasizes the application of general principles whose proofs are purposely omitted. Links the subject areas of economics and operations research. Many good illustrations of solutions to managerial planning problems (e.g., production decision, transport planning, location of activities). No exercises. A handy supplement book. I-CH

APPLICATIONS (PHYSICS), T(16-18), L. *Supersonic Flow and Shock Waves*. R. Courant, K.O. Friedrichs. Appl. Math. Sci., V. 21. Springer-Verlag, 1976, xvi + 464 pp, $19.80. Unaltered reprint of a 1948 Interscience publication, itself based on a 1944 OSR report. A classic treatise on nonlinear waves and concommitant shock fronts. LAS

APPLICATIONS (PHYSICS), T(15-16: 1), S, L*. *Celestial Mechanics*. Harry Pollard. Carus Math. Mono., No. 18. MAA, 1976, x + 134 pp, $11. A corrected reprinting of Chapters I-III of the author's 1966 *Mathematical Introduction to Celestial Mechanics* (TR, February 1967): central force problem, n-body problem, Hamilton-Jacobi theory. A concise, elegant primer. LAS

APPLICATIONS (PHYSICS), T(18: 2, 3), S, P. *Quantum Theory of Open Systems*. E.B. Davies. Acad Pr, 1976, x + 171 pp, $16.50. Classical quantum mechanics is formulated for closed systems which do not interact with the outside world. The present monograph provides a mathematical foundation (with functional analysis on Hilbert space as a prerequisite) for quantum theory of open systems. Applications are in quantum optics and irreversible dynamics. PJM

APPLICATIONS (PHYSICS), T(18: 1), P. *Variational Methods in Theoretical Mechanics*. J.T. Oden, J.N. Reddy. Springer-Verlag, 1976, x + 302 pp, $14.80 (P). A well-written text for a post-functional analysis course for students of mechanics and engineering science, but of value to all interested in modern variational methods. 148 references; no index. RBK

APPLICATIONS (PHYSICS), S(11-14), L*. *Mathematical Methods in Science*. George Pólya. MAA, 1977, xi + 234 pp, $4.50 (P). A corrected and slightly revised republication of the 1963 SMSG Studies in Mathematics, V. XI, itself a supplement to *Mathematics and Plausible Reasoning*. Astronomy, statics, dynamics motivate the mathematics of approximation, vectors, differential equations, all in the author's distinctive heuristic style. LAS

APPLICATIONS (PHYSICS), P. *Mathematical Physics and Physical Mathematics*. Ed: Krzysztof Maurin, Ryszard Raczka. Reidel, 1976, xviii + 504 pp, $39. Proceedings of the March 1974 International Symposium in Warsaw organized by the Mathematical Institute of the Polish Academy of Sciences, the Institute for Nuclear Research and the University of Warsaw. JAS

APPLICATIONS (PHYSICS), P. *Lecture Notes in Mathematics-559: Le Mouvement Brownien Relativiste*. Jean-Pierre Caubet. Springer-Verlag, 1976, ix + 212 pp, $10.20 (P).

APPLICATIONS (SOCIAL SCIENCE), T(17-18: 1). *Fundamentals of Decision Theory*. D.J. White. North-Holland, 1976, xiii + 387 pp, $35. Intended for mathematically-inclined economics students. The concept of decision is not confined to the statistical area. "The essential emphasis is on how one may formally state what it is one knows about a decision situation, including the preference characteristics of [the decision makers], in order to make a decision." Little specific mathematics assumed, as the author is not concerned with the mathematics of arriving at solutions; but the reader must be comfortable with abstraction in structures and concepts. Many references to an extensive bibliography. PJC

Reviewers Whose Initials Appear Above

Richard J. Allen, St. Olaf; John Dyer-Bennet, Carleton; Ceceila Bleecker, Carleton; Paul J. Campbell, St. Olaf; Clifton E. Corzatt, St. Olaf; Steven Galovich, Carleton College; Ih-Ching Hsu, St. Olaf; Lorraine L. Keller, St. Olaf; Roger B. Kirchner, Carleton; Richard S. Kleber, St. Olaf; Loren C. Larson, St. Olaf; Pierre J. Malraison, Carleton; R.W. Nau, Carleton; J. Arthur Seebach, St. Olaf; Lynn Arthur Steen, St. Olaf.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, NW, Washington, D. C. 20036. Items must be submitted at least five months before publication can take place.

PERSONAL ITEMS

Texas Tech University: Dr. William Gustafson, University of Indiana, has been appointed Assistant Professor; Dr. David Lutzer, University of Pittsburgh, has been appointed Visiting Professor; Dr. Barry Turett, University of Illinois, has been appointed Lecturer; Associate Professors Shelby Hildebrand and Horace Woodward have been promoted to Professors.

University of Akron: Dr. T. E. Price, Jr., University of Georgia, has been appointed Assistant Professor; Associate Professor Louis Ross has been promoted to Professor.

University of Texas, Arlington: Associate Professors A. R. Mitchell and R. W. Mitchell have been promoted to Professors; Assistant Professor J. T. B. Beard, Jr., has been promoted to Associate Professor; Assistant Professor G. B. Turney has been promoted to Professor.

Dr. E. D. Fife, Wesleyan University, has been appointed Assistant Professor at Mary Washington College.

Dr. S. A. Levin, Chairman of the Section of Ecology and Systematics at Cornell University since 1974, has been appointed Professor of Ecology and Applied Mathematics in the Division of Biological Sciences, Cornell University.

Assistant Professor P. J. Murray, Chairman of the Mathematics Department at St. Martin's College, has been promoted to Associate Professor.

Mr. Edward J. Bandur, Long Beach, California, died on December 6, 1976. He was a member of the Association for eleven years.

Professor Harry N. Carter, University of Tulsa, died on June 22, 1976, at the age of 64. He was a member of the Association for thirty years.

Mr. Hubert C. Dixon, Chairman of the Department of Mathematics and Physics at Gardner-Webb College, died on April 16, 1976, at the age of 64. He was a member of the Association for seven years.

Miss Jane L. Evans, St. Petersburg Junior College, died on October 30, 1976. She was a member of the Association for seventeen years.

Professor Emeritus Olan H. Hamilton, Oklahoma State University, died on August 31, 1976, at the age of 77. He was a member of the Association for thirty-eight years.

Dr. Ben Zion Linfield, University of Virginia, died on August 22, 1976, at the age of 79. He was a member of the Association for fifty-three years.

Professor Emeritus Wilson L. Miser, Vanderbilt University, died on February 28, 1974, at the age of 88. He was a Charter Member of the Association.

Associate Professor Donald K. Pease, University of Connecticut, died on December 3, 1975, at the age of 60. He was a member of the Association for thirty-six years.

Professor Albert Soglin, City College of Chicago, Loop College, died on October 4, 1976, at the age of 58. He was a member of the Association for twenty-eight years.

Dr. William Harold Wilson, University of Florida, died on November 11, 1976, at the age of 83. He was a member of the Association for fifty-nine years.

MODULES IN APPLIED MATHEMATICS

The MAA sponsored a College Faculty Workshop entitled Modules in Applied Mathematics which was held at Cornell University from July 26 to August 20, 1976. This activity was supported by the National Science Foundation through grant No.HES 75–00713. Sixty educational modules were produced in connection with this Workshop, and these are listed in the two-page house ad at the end of this issue, along with a brief indication of prerequisites. These modules are available free of charge while the initial supply lasts. Due to the limited quantities, however, orders must be restricted to eight modules and no more than 300 total pages. Copies may be obtained from the Workshop Director, W. F. Lucas, at the following address: MAA Workshop, 334 Upson Hall, Cornell University, Ithaca, N.Y. 14853.

PROGRAM OF INSTRUCTIONAL LECTURES AT THE JOHNS HOPKINS UNIVERSITY

A program of instructional lectures on "Applied Matrix Computations" will be given by Professor Gene H. Golub, Stanford University, at The Johns Hopkins University from August 15 to 19, 1977. The program is sponsored by The Johns Hopkins Mathematical Sciences Department and the Johns Hopkins Press.

In a series of lectures given in the mornings throughout the week Professor Golub will provide background material and present computational methods. In the afternoons Professor Golub and guest lecturers will discuss example applications and computer implementations. Applications to problems in statistics, differential equations and optimization will be stressed.

Time will be held open later each afternoon for contributed papers on topics in the broad area of computational linear algebra and its applications.

Further information may be obtained by writing to Dr. Richard Bartels, Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, Maryland 21218.

WASHINGTON STATE UNIVERSITY–MATHEMATICS CAREER REENTRY FOR WOMEN

The Washington State University Program *Mathematics Career Reentry for Women* is a project funded by the National Science Foundation in an effort to tap the underutilized mathematical resource which women represent.

The Career Reentry program falls naturally into two parts — a five week short course beginning August 8, 1977, and a year of formal study.

For further information, write to Professor Calvin T. Long, Department of Pure and Applied Mathematics, Washington State University, Pullman, Washington 99164.

CONFERENCE AT THE EVERGREEN STATE COLLEGE

The Evergreen State College with support from the National Science Foundation is sponsoring a conference on "Self-Paced Learning, The Concept and Process." The conference will take place during the weeks of August 1–5 and 8–12, 1977. Subject areas will be mathematics, natural and social science. For more information write Frederick D. Tabbutt, The Evergreen State College, Olympia, Washington 98505.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

OCTOBER MEETING OF THE SOUTHERN CALIFORNIA SECTION

The second annual joint meeting of the Southern California Sections of SIAM, MAA and ASA was held at the Jet Propulsion Laboratory at Pasadena, California, on October 28, 1976, with a registered attendance in excess of 280. Dr. X. X. (Skip) Newhall of the Jet Propulsion Laboratory was the Program Chairman, and welcomed the attendees of the three sections.

There were three sessions and a tour of JPL. The sessions were chaired by Alfred Inselberg, IBM, Chairman, Southern California Section, SIAM, Nancy Minter, Office of the Mayor, Los Angeles, Chairman, Southern California Section, ASA, and Paul Yale, Pomona College, Chairman, Southern California Section, MAA.

The following program was presented:

Recent developments in earthquake prediction research, by J. H. Whitcomb, California Institute of Technology.

Catastrophes: An overview of theory and applications, by Craig Benham, Lawrence University and California Institute of Technology.

Panel Discussion: How flexible is today's mathematics degree? moderated by M. L. Juncosa, RAND Corporation. The panelists were D. L. Bentley, Pomona College, Eldon Hansen, Lockheed Research Laboratories, C. L. Lawson, JPL, and E. M. Scheuer, California State University, Northridge.

Mathematical theory of stock options, by Edward Thorpe, University of California, Irvine.

Aspects of randomized-response survey techniques, by William Barksdale, McDonnell-Douglas Corporation.

Computer processing of Viking Lander Imagery, by W. B. Green, JPL.

E. I. DEATON, *Secretary-Treasurer*

CALENDAR OF FUTURE MEETINGS

Fifty-seventh Summer Meeting, University of Washington, August 14–16, 1977.

Sixty-first Annual Meeting, Atlanta, Georgia, January 6–8, 1978.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN, last weekend in April or first weekend in May. Deadline for papers 6 wks. bef. mtg.

FLORIDA, early March. Deadline for paper titles 2 wks. bef. mtg.

ILLINOIS, first Friday/Saturday in May.

INDIANA

INTERMOUNTAIN

IOWA, third weekend in April. Deadline for papers February 1.

KANSAS, March or April. Deadline for papers January 1.

KENTUCKY, early April. Deadline for papers 6 wks. bef. mtg.

LOUISIANA-MISSISSIPPI, Buena Vista Hotel-Motel, Biloxi, Mississippi, February 17–18, 1978.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA, Saturday before Thanksgiving and last Saturday in April.

METROPOLITAN NEW YORK, Spring. Deadline for papers 2 wks. bef. mtg.

MICHIGAN, first Friday and Saturday in May. Deadline for papers 6 wks. bef. mtg.

MISSOURI, late March/early April. Deadline for papers January 31.

NEBRASKA, April.

NEW JERSEY, early November and early May.

NORTH CENTRAL, University of Minnesota, Morris, October 14–15, 1977.

NORTHEASTERN, Middlebury College, Middlebury, Vermont, June 18–19, 1977.

NORTHERN CALIFORNIA, College of Notre Dame, Belmont, February 1978.

OHIO

OKLAHOMA-ARKANSAS , (approx.) Friday and Saturday of first weekend in April. Deadline for papers 3 wks. bef. mtg.

PACIFIC NORTHWEST, second Saturday in June. Deadline for papers 6 wks. bef. mtg.

PHILADELPHIA, Moravian College, Bethlehem, November 19, 1977.

ROCKY MOUNTAIN, last weekend in April or first in May. Deadline for papers 8 wks. bef. mtg.

SEAWAY, first Saturday in November and Saturday in late April. Deadline for papers 6 wks. bef. mtg.

SOUTHEASTERN, Clemson University, Clemson, South Carolina, Spring 1978.

SOUTHERN CALIFORNIA, California State Polytechnic University, San Luis Obispo, November 11–12, 1977.

SOUTHWESTERN, usually in April. Deadline for papers 2 wks. bef. mtg.

TEXAS, Friday and Saturday in early April. Deadline for papers March 1.

WISCONSIN, Friday and Saturday between mid-April and first week in May. Deadline for papers 6 wks. bef. mtg.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Washington, February 12–17, 1978.

AMERICAN MATHEMATICAL ASSOCIATION OF TWO YEAR COLLEGES, Peachtree Plaza Hotel, Atlanta, October 13–14, 1977.

AMERICAN MATHEMATICAL SOCIETY, University of Washington, August 15–18, 1977.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION, University of North Dakota, Grand Forks, June 13–16, 1977.

ASSOCIATION FOR COMPUTING MACHINERY, Olympic Hotel, Seattle, Washington, October 17–19, 1977.

ASSOCIATION FOR SYMBOLIC LOGIC, Wrocław, Poland, August 1–2, 1977.

ASSOCIATION FOR WOMEN IN MATHEMATICS

CANADIAN SOCIETY FOR HISTORY AND PHILOSOPHY OF MATHEMATICS/SOCIÉTÉ CANADIENNE D'HISTOIRE ET DE PHILOSOPHIE DES MATHÉMATIQUES, Hamilton, Ontario, June 2, 1977.

FIBONACCI ASSOCIATION

INSTITUTE OF MATHEMATICAL STATISTICS, Seattle, Washington, August 14–18, 1977.

MU ALPHA THETA, Loras College, Dubuque, Iowa, August 7–10, 1977.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, San Diego, California, April 12–15, 1978.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Peachtree Plaza Hotel, Atlanta, November 7–9, 1977.

PI MU EPSILON, University of Washington, August 14–16, 1977.

SCHOOL SCIENCE AND MATHEMATICS ASSOCIATION, William Penn Hotel, Pittsburgh, Pennsylvania, November 10–12, 1977.

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, Philadelphia-Sheraton Hotel, Philadelphia, June 13–15, 1977 (25th Anniversary Meeting).

MODULES IN APPLIED MATHEMATICS (See the **NOTICE** on page 506 of this issue.)

Code	Author	Title	Pages	Prerequisites
Ba 1	Baker, Robert L., Jr.	Five Nautical Models	17	hs, c
Ba 2	Baker, Robert L., Jr.	Car Following Models	32	c, de
Bo 3	Bolger, Edward M.	Proportional Representation	19	a
Br 4	Borrelli, Robert L.	Shaking a String to Rest	26	c
Br 5	Borrelli, Robert L.	Stability of a Tennis Racket	38	c, la, de
Bs 6	Brams, Steven J.	One Man, *n* Votes	19	hs, st
Bn 7	Braun, Martin	The Van Meegeren Art Forgeries	10	de
Bn 8	Braun, Martin	Single Species Population Model	12	de
Bn 9	Braun, Martin	The Spread of Technological Innovations	9	de
Bn10	Braun, Martin	A Model for the Detection of Diabetes	10	de
Bn11	Braun, Martin	Why the Percentage of Sharks Caught in the Mediterranean Sea Rose Dramatically during World War I	10	de
Bn12	Braun, Martin	The Principle of Competitive Exclusion in Population Biology	9	c, de
Bn13	Braun, Martin	A Model for the Spread of Gonorrhea	13	c, de
Cb14	Cobb, Loren	Stochastic Difference Equations with Sociological Applications	30	a, p or s
Co15	Coleman, Courtney	Quadratic Population Models: Almost Never Any Cycles	19	c, de
Co16	Coleman, Courtney	Combat Models	22	de
Co17	Coleman, Courtney	Hilbert's 16th Problem—How Many Cycles?	23	de
Co18	Coleman, Courtney	Biological Cycles and the Five-Fold Way	31	a, de
Da19	Davis, Morton	Machine Learning of Games	23	hs
DP20	Deegan, John, Jr. and Packel, Edward W.	To the (Minimal Coalition) Victors Go the (Equally Divided) Spoils: a New Power Index for Simple *n*-Person Games	17	a
Dr21	Drew, Donald A.	Equilibrium Speed Distributions	14	c, p
Dr22	Drew, Donald A.	How Long Should a Traffic Light Remain Amber?	7	c
Dr23	Drew, Donald A.	Queue Length at a Traffic Light via Flow Theory	10	c
Dr24	Drew, Donald A.	Traffic Flow Theory	17	c, de
Dr25	Drew, Donald A.	Surge Tank Analysis	14	c, de
Dr26	Drew, Donald A.	Cigarette Filtration	22	c, de
Fe27	Fennell, Robert E.	Population Growth—An Age Structure Model	15	la, cp
Fr28	Frauenthal, James C.	Difference and Differential Equation Population Growth Models	24	c
FS29	Frauenthal, James C. and Saaty, Thomas L.	Foresight—Insight—Hindsight	23	none
G 30	Greenspan, Donald	An Arithmetic Model of Gravity	18	hs
He31	Heaney, James P.	Urban Wastewater Management Planning	13	e, gt
Hn32	Henderson, Beverly	Qualitative Solution Sketching for First Order Ordinary Differential Equations	24	c
Hn33	Henderson, Beverly	Setting Up First Order Differential Equations from Word Problems	31	c
LB34	Lucas, William F. and Billera, Louis J.	Modelling Coalition Values	40	hs, st

(Continued from previous page.)

MODULES IN APPLIED MATHEMATICS

(See the **NOTICE** on page 506 of this issue.)

Code	*Author*	*Title*	*Pages*	*Prerequisites*
Ma35	Maceli, John C.	How to Ask Sensitive Questions Without Getting Punched in the Nose	12	p
Mr36	Marcus-Roberts, Helen	DNA, RNA and Random Mating—Some Simple Applications of the Multiplication Rule	27	ar
Mr37	Marcus-Roberts, Helen	A Comparison of Some Stochastic and Deterministic Models of Population Growth	59	c, de, p
MR38	Marcus-Roberts, Helen and Roberts, Fred S.	Malaria (Models of the Population Dynamics of the Malaria Parasite)	21	p
Mo39	Marrero, Oswaldo	A Model for an Epidemic of a Contagious Disease	12	c
Mo40	Marrero, Oswaldo	An Optimal Inventory Policy Model	8	c
My41	Mayer, Lawrence S.	An Analysis of Alternative Voter Registration Systems	63	c, s
Pa42	Packel, Edward W.	Four-Way Stop or Traffic Light? An Illustration of the Modelling Process	20	cp, c
Py43	Perry, E. L.	A Pulse Process Model of Athletic Financing	33	ma
Pt44	Peterson, Elmor L.	Traffic Equilibria on a Roadway Network	60	c, la, lp
Pr45	Prather, Ronald E.	Finite Covering Problems	31	st, a
Ri46	Rice, Peter	Committee Decision Making	20	hs
Ro47	Roberts, Fred S.	Efficiency of Energy Use in Obtaining Food I: Humans	44	lp
RM48	Roberts, Fred S. and Marcus-Roberts, Helen	Efficiency of Energy Use in Obtaining Food II: Animals	79	c
SD49	Saaty, Thomas L. and Diaz-Mora, Ruben	A Guided Tour Through Graph Theory Algorithms	180	a, ma
Sa50	Saaty, Thomas L.	Hierarchies, Reciprocal Matrices and Ratio Scales	56	la
So51	Solomon, Daniel L.	The Spatial Distribution of Cabbage Butterfly Eggs	20	c, p
So52	Solomon, Daniel L.	Conditional Expectation and Caviar	7	c, p
St53	Straffin, Philip D., Jr.	Power Indices in Politics	86	a, p
TP54	Thrall, Robert M. and Perry, E. L.	An Everyday Approach to Matrix Operations	29	ar
To55	Todd, Michael J.	Computing Fixed Points with Applications to Economic Equilibrium Models	55	la, lp, e
Tu56	Tuchinsky, Philip M.	Least Squares, Fish Ecology and the Chain Rule	56	c, de
U 57	Uslaner, Eric M.	Vote Trading in Legislative Bodies: Opportunities, Pitfalls and Paradoxes	39	st
W 58	Weber, Robert J.	Multiple-Choice Testing	33	ma, c, p
Z 59	Zahavi, Jacob	Power Systems Performance in Face of Supply and Demand Uncertainty	55	c, p
Z 60	Zahavi, Jacob	An Optimal Mix Problem	14	c

PREREQUISITES:

ar = arithmetic	st = set theory	de = differential equations	s = statistics
a = elementary algebra	ma = matrix algebra	lp = linear programming	e = some economics
hs = high school math	la = linear algebra	gt = game theory	
cp = computer programming	c = calculus	p = probability	

"The finest book ever written on the subject."

Steven T. Hedatniemi
Associate Professor
of Computer Science
UNIVERSITY OF VA

For a complimentary copy, please send course title and approximate enrollment to:
James A. Ryder
HOLT, RINEHART AND WINSTON
383 MADISON AVENUE
NEW YORK, NY 10017

COMBINATIONAL OPTIMIZATION: NETWORKS AND MATROIDS

EUGENE L. LAWLER, University of California, Berkeley

This up-to-date and well-written text has been hailed as the landmark survey of network flow theory.

Emphasizing problem formulations and applications, Lawler uses a theoretical development for network flows, matching, and matroid intersection problems.

The problems employ augmentation of feasible solutions ... means of a labeling procedure ... and linear programming duality.

Although intended for an operations research audience, the text takes a modern computer science viewpoint. Special features include:

- The first textbook discussion ever of non-bipartite matching and matroid optimization
- A chapter on mathematical preliminaries
- Algorithms stated in English
- Many theoretical problems and exercises to encourage research

Elements of Set Theory

By HERBERT B. ENDERTON

Set theory is important to contemporary mathematics, and the student will find this book enlightening not only for its treatment of set theory, but for the applications of set-theoretic concepts to other parts of mathematics and to other disciplines. The author's axiomatic development (based on the Zermelo-Fraenkel axioms) takes the student step by step through set theory and fully discusses cardinal and ordinal numbers. The extensive explanations as well as the range of exercises will assist the reader in maneuvering with ease through the material. The author admirably fulfills his goal of presenting a modern viewpoint on this subject in a readable and cogent manner. A major advantage of this work is its flexibility; axiomatic material is specially indicated to enable the user to deemphasize it or to omit it altogether. Explanations and motivating discussions as well as numerous examples, historical comments, and diagrams, not only heighten the ease with which students may grasp this material, but also increase interest in the subject.

Because of its great range, this work will be valuable to the beginner, the intermediate, and, in its last chapters, the sophisticated student in this field. Students and teachers of set theory as well as other related phases of mathematics will find this book of enormous value. The book will lend itself to use in undergraduate courses in set theory, and will be a welcome and stimulating supplementary textbook for mathematicians, mathematics students, philosophy students, and all those needing a background in set theory.

1977, 288 pp., \$12.95/£9.25
ISBN: 0-12-238440-7

For complimentary copies, write to the Sales Department, Academic Press, New York. Please indicate course, enrollment, and present textbook.

Send payment with order and save postage plus 50¢ handling charge. *Prices are subject to change without notice.*

ACADEMIC PRESS, INC.
A Subsidiary of
Harcourt Brace Jovanovich, Publishers
111 FIFTH AVENUE, N.Y., N.Y. 10003
24-28 OVAL ROAD, LONDON NW1 7DX

THE MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.
Washington; DC 20036

THE AMERICAN
MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

AN OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 84 **NUMBER 7**

CODEN: AMMYAE

CONTENTS

(*Continued on inside cover*)

AUGUST-SEPTEMBER **1977**

NOTICE TO AUTHORS

Research papers per se are unsuitable; see statement of policy (Vol. 81, p. 1).

Please follow the format in current issues of the Monthly. Manuscripts must be legibly typewritten or reproduced from typewritten copy, double spaced with wide margins, and on one side of the paper. Three copies should be submitted to the appropriate editor and one kept by the author as protection against loss. The author's full address *must* appear at the end of the manuscript.

Backlog: Main Articles 21 months, Math. Notes 18 months, Research Problems 17 months, Classroom Notes 16 months, Math. Education 17 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to R. P. Boas, Department of Mathematics, Northwestern University, Evanston, IL 60201; NOTES, etc.: to the corresponding Associate Editor; REPRINT PERMISSION: to Leonard Gillman, Mathematical Association of America, University of Texas, Austin, Texas 78712 (see also the copyright notice below); ADVERTISING CORRESPONDENCE: to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo, Buffalo, N.Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. Willcox, Mathematical Association of America, 1225 Connecticut Ave., N. W., Washington, D.C. 20036; BACK ISSUES: Contact P. and H. Bliss Co., Middletown, CT 06457.

Annual dues for members of the Association, including a subscription to the American Mathematical Monthly, are $18.00 for each of the first two years of membership and $22.00 thereafter. Student Membership is available with annual dues of $12.00. For nonmembers the subscription price is $25.00.

Published by the Association at Washington, D.C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D.C., and additional mailing offices.

ARRANGE YOUR OWN SABBATICAL EXCHANGE: USE SEIS

Do you feel the need for a change of scene for a year or a semester, but are not eligible for a sabbatical leave? Why not investigate the possibility of arranging a "no-cost sabbatical" by exchanging positions with a faculty member with similar interests on another campus? The stimulation of a sabbatical exchange can benefit both you and your institution and cost little or nothing. *See the details of the MAA Sabbatical Exchange Information Service on page* 590 *of this issue.*

IMPRESSIONS OF MATHEMATICAL EDUCATION IN THE PEOPLE'S REPUBLIC OF CHINA

VICTOR KLEE

Introduction. Perhaps the title should read "Fleeting Impressions...," for the delegation spent less than a month in China and most of us did not speak Chinese. Our aim was to learn about mathematical research, applications of mathematics, and some aspects of mathematical education. We visited a research institute, several universities, and a lot of factories, but only one secondary school and no primary school. In the case of most countries, it would be ludicrous to write an article on the basis of so little information. However, China is such an important country, its current mathematical scene is so little known outside China, and our colleagues have shown so much interest in our trip, that it has seemed advisable to write this account. What follows is an attempt to share our experiences (and our ignorance!), insofar as they relate to mathematical education in China.

The Pure and Applied Mathematics Delegation, sent to China in May of 1976, consisted of nine mathematicians — Saunders Mac Lane of the University of Chicago (Chairman), George Carrier of Harvard (Vice-Chairman), Edgar Brown of Brandeis, Walter Feit of Yale, Joseph Keller of the Courant Institute, Victor Klee of the University of Washington, Joseph Kohn of Princeton, Henry Pollak of the Bell Telephone Laboratories, Hung-hsi Wu of the University of California at Berkeley — and also Carl Leban, an orientalist from the University of Kansas, and Anne Fitzgerald, a staff member of the organization sponsoring the trip. That organization, the Committee on Scholarly Communication with the People's Republic of China (CSCPRC), is supported by the National Academy of Sciences, the Social Science Research Council, and the American Council of Learned Societies. The CSCPRC handles the US side of the cultural, scientific and technological exchanges that have resulted from the 1972 Shanghai Communiqué. By the end of 1976 the CSCPRC had sent 19 delegations to China in a variety of fields and had hosted 25 Chinese delegations. Although this was the first exchange in mathematics sponsored by the CSCPRC, several American mathematicians and computer scientists have visited China independently since 1972. Through these visits a foundation is being laid for continuing scientific and cultural interaction between Americans and Chinese. The Mathematics Delegation hopes that its visit will be followed by continuing communication with Chinese mathematicians.

In Peking, all members of the delegation visited Peking University, Tsinghua University, the Institute of Mathematics of the Chinese Academy of Science, and a number of factories. Those with more applied interests then went north to Harbin, to learn about industrial, agricultural and forestry operations in Heilungkiang Province and to confer with faculty members from Heilungkiang University. Then all went south to Soochow for a day's sightseeing and on to Shanghai for visits to Futan University and the Shanghai Hua-tung Normal University. Several members visited the May 7th Middle School and spent an evening with part of a group that plans the mathematics curriculum for primary and secondary schools in Shanghai.

Most of this article consists of excerpts from a more detailed report prepared by the entire delegation. To supply the necessary historical background, excerpts from a book by Frank Swetz have

also been included. These publications are listed at the end of the article, along with other references that may interest the reader.

Historical background. Mathematics is probably viewed by most of its Western practitioners as being relatively acultural and apolitical. However, mathematical education in China has been greatly influenced by cultural and political changes. The purpose of this section is to describe the historical background for recent developments. The next six paragraphs are excerpted from the excellent book by Swetz, mainly pp. 51–52, 102–105, 115, 135, 154, 205–207.

There was significant mathematical activity in China at a very remote date, and some of it anticipated similar work in other parts of the world by several centuries. Important Chinese mathematical discoveries were made as late as the fourteenth century A.D., but in the ensuing six centuries the development of Chinese mathematics was very slow, perhaps because of a social and moral code that did not encourage the applications of science and because of religious doctrines that discouraged scientific conjecture.

With the impact of the West in the early nineteenth century, China, lacking modern industry and armaments, was at the mercy of foreign invaders. In their search for remedies, some Chinese officials recognized the need for science-oriented education and campaigned for the addition of mathematics to the classical curriculum. After a long struggle, mathematics was again accepted as a subject worthy of study (it first appeared on the Traditional Civil Service Examination in 1888) but broad mathematics education reforms were difficult to achieve. China entered the twentieth century committed to the establishment of a national school system with a modern curriculum but the lingering prejudices of her traditional past would impede this transformation.

The first half of the twentieth century was a period of great educational flux in China. A modern school system had been established, but traditional educational thinking still dominated the school scene. Methods of instruction encouraged rote memorization and provided little opportunity for individual inquiry. Mathematical studies bore little relevance to the requirements of an agrarian society or industrial application.

In 1947, near the end of the Nationalist era, the Ministry of Education suggested that mathematics instruction should provide China's young with the mathematical skills they needed to make a living. During their Yenan exile (1935–49) the Communists developed similarly utilitarian ideas based on "proletarian relevance." They especially wanted to eliminate "elitism" on the part of the teachers and "championism" among the students, and to turn classroom work from an individual into a collective endeavor. However, during their first eight years in power (1949–57) they were unable to effect any broad reforms in these matters. Two important developments of those years were, on the one hand, extensive adoption of Soviet models of learning and methods of instruction and, on the other hand, the nationalization of private universities which paved the way for ultimate rejection of all foreign influences in education.

In 1958–60, as part of the Great Leap Forward, Soviet assistance was rejected and the Chinese sought once again to develop an educational model consistent with worker-peasant needs. A variety of proletarian-oriented educational experiments were attempted. Regular school curricula, including mathematics studies, were drastically abbreviated, and labor requirements took a toll of student time and energy. These excesses lowered mathematical standards to a level that warranted national concern, and resulted in a return to a rigorous standardized curriculum. By the end of 1962, education in China had reached an "elitist" extreme with the establishment of a system of privileged schools for talented students; some of these schools devoted as much as half of their class time to mathematics and science. This accommodated the state's scientific goals, but basic socialist tenets were violated by the student selection process. Led by Mao, the Communist Party began open attacks on the school system, and finally, in May of 1966, Mao called upon the students to overthrow the educational system. They responded and the Great Proletarian Cultural Revolution swept over China. Many experienced and skilled teachers who had received Western training were publicly humiliated as bourgeois revisionists, and forced to confess their crimes. Complete disruption of formal education forced schools to close in the summer of 1966.

After a prolonged closing, the reopened schools revealed a completely restructured system based on "proletarian relevance." Mathematics was still given high priority, but the new mathematics courses were narrower in scope and lower in quality than those before the Cultural Revolution. The inclusion of political materials in mathematical studies was greatly increased. Mathematics was designed for immediate societal needs rather than as preparation for higher learning. All study had become a collective endeavor and tests were abolished. (This ends the excerpts from Swetz.)

The twin shackles of tradition and foreign influence having been removed, the Chinese educational system could truly serve the needs of the people. But would it? Some Chinese leaders claimed that being expert was more important than being "red," that the schools no longer produced experts, and that the fulfillment of China's scientific and technical needs was suffering in consequence. In short, they felt that the failings of the Great Leap Forward were being repeated. An articulate spokesman for this view was Teng Hsiao-p'ing, who once suggested that the color of a cat was less important than its mousecatching ability. Teng had been purged during the Cultural Revolution but then returned to power, and as late as our pre-trip briefing was widely regarded as the probable successor of Chou En-Lai as premier. However, shortly before our visit the famous demonstration in Peking's Tien An Mien square occurred and Teng was again purged, being stripped of all his Party posts.

At each place we visited in China, the official welcoming speech dwelt on the evils of Teng Hsiao-p'ing, the "unrepentant capitalist roader" who wanted to "reverse the correct historical verdicts of the Cultural Revolution." But more recently, after the death of Mao Tse-tung and the propaganda campaign against the "gang of four", there seems to have been a movement to rehabilitate Teng to some extent. Does this foretell another change in the Chinese educational system?

Mathematical education in the secondary schools. Pre-university education in China usually consists of five years in a primary school followed by five years in a secondary school. The secondary schools are called "middle schools."

At the May 7th Middle School in Shanghai, the Principal said "The purpose of our school is to turn all the students into ordinary workers and peasants." When asked about "dropouts," she said "It has never happened. All students complete the course." (The course then lasted four years, but is being raised to five.) All students follow the same curriculum, which involves 28 45-minute classes per week. Each teacher has 10 classes per week. For each of the students' four years, 4 or 5 of the 28 classes are in mathematics. For fourth-year students there are 5 mathematics, 2 basic industrial knowledge (including physics and chemistry), 2 agriculture (including biology), 5 Chinese, 4 English, 2 geography, 1 history, 2 literature and revolutionary arts (including drawing), 2 physical training. Examinations are given, but they are less formal and more practical than those before the Cultural Revolution, "when the student took the teacher as his enemy."

We spent an evening with five members of the 47-member team that plans the mathematics curriculum for most of the primary and middle schools of Shanghai. The team has been together for nine years, includes 12 women and three workers, and its members range in age from 20 to 71. The team communicates with similar ones in other regions by exchanging materials and personnel. Within broad guidelines there is some local autonomy in developing the mathematics curriculum, but most curricula are similar to Shanghai's. Much of the local variation stems from Chairman Mao's advice to "combine theory with practice," for the "practice" included in the curriculum depends on the needs of local industry or agriculture.

Of the five members of the team, one had been a middle school mathematics teacher, two were mathematics graduates of Shanghai Normal College, one was a chemistry graduate of Futan University, and one had been a worker (and member of the propaganda team) in a factory making electrical machinery. Their carefully prepared presentation began with a review of the evils of the "old system," in which algebra, plane and solid geometry, trigonometry, and analytic geometry were taught as separate subjects, all on a theoretical basis. There was no attempt to explain the practical importance of the material. "Many graduates didn't like to combine with workers and peasants" and "Chairman Mao's revolutionary line couldn't be implemented." However, after "the Great Pro-

letarian Cultural Revolution destroyed the revisionist line," it was possible to modify education so as to take account of the "three great revolutionary movements": class struggle, the revolution in production, and the revolution in scientific experiment. One of the goals of Chinese education today is to teach the students to "love socialist society, and to love the workers and peasants."

We were told that their current approach to mathematics is based explicitly on certain aspects of dialectical materialism, especially on the belief that learning proceeds "from practice to general knowledge and again to practice." Engels was quoted a number of times, in particular as saying that "mathematics is supplementary to dialectical materialism, and expresses a form of dialectical materialism." This was illustrated by the "dialectic contradiction" between positive and negative numbers, powers and roots, constants and variables, and differentials and integrals, all of which are "pairs of opposites that depend on each other." So much time was spent in presenting the political and philosophical background of the mathematics curriculum that there was not much left for discussing details of the subject matter. The curriculum is still experimental, and most texts are rewritten every year. Despite our requests, it was impossible for us to see copies of the texts. Nevertheless, the following details emerged from the discussion:

(1) Mathematics from the first through the tenth year (that is, through the five years of primary school and the five years of middle school), is taught as a unified course, with no artificial divisions between subjects and with much attention to practical applications. The study is supposed to be guided by dialectical materialism, by the desires to "combine form and number" and to "combine theory with practice."

(2) The first two years of middle school are devoted to basic facts from algebra and geometry. The Pythagorean theorem is proved by area considerations based on subdividing a square. Parallel lines are introduced intuitively in terms of the equality of corresponding angles when cut by a transversal, and equality of the alternate interior angles is then "proved." Similar plane figures are introduced as a basis of measurement and are then applied to actual measurements in surveying, building of bridges and ships, and so on. In general, the practical utility of various geometric figures and constructions (e.g., finding the center of a circle or a circular arc through three given points) is emphasized. The pre-Cultural Revolution axiomatic approach to plane geometry has been largely abandoned, though some aspects of formal logic (as a means of deducing useful properties from other useful properties) are retained in an unsystematic way.

(3) Little was said about algebra. It was mentioned that the irrationality of $\sqrt{2}$ had been proved in the past, but not at present. One project, carried out in conjunction with a factory, was to write a program for computer control of a milling machine, and this involved some use of boolean algebra and expression of numbers in the binary system. Expression in other bases is also taught. Complex numbers are apparently not mentioned in the middle school.

(4) In plane analytic geometry, the discussion of polar coordinates includes the equations of Archimedean spirals and their uses in cam designs. The notion of parametric equations of a curve is introduced by means of the cycloid and its evolute, and their uses in various production problems are explained.

(5) The third and fourth year of middle school mathematics are devoted to logarithms, the function concept, and basic concepts of calculus. Natural logarithms are mentioned, but their importance is not explained to the student. Differentiation and integration of polynomials are based on geometric intuition. The notion of derivative is motivated by considering the tangent line as a limit of secants.

(6) In the fifth (last) year of middle school, the students learn more calculus, some statistics, and some elementary methods of operations research. These include the use of orthogonal experimental designs, and the use of golden-section search for optimization of a unimodal function of one variable.

(7) Considerable attention is devoted to computation. During their work in factories, students have access to a computer. Some schools have built their own computers.

(8) All mathematics training is intended to reflect the belief that "although mathematics is abstract in appearance, it originated from the real world and is a product of the needs of the people." Thus, in

addition to frequently hearing about practical problems in the classroom, each student spends a month during the school year on a farm or at a factory. There the student is supposed to be taught about applications of mathematics and is encouraged to look for more applications.

Mathematical education in the universities. In each university visit, an introductory speech was given by a vice-chairman or "responsible member" of the Revolutionary Committee. It invariably began by describing the local campaign against Teng Hsiao-p'ing and the struggle to combat the "right deviationist wind." Next came a brief description of the institution, contrasting its present status with that before the Cultural Revolution. The words did not vary greatly from one place to another. We were told that in the old system, students crammed for examinations (an analogy was drawn with the forced feeding of Peking duck), then threw away their notebooks and forgot everything after the examinations. Teachers considered themselves superior to students, and students held the workers and peasants in contempt. Now, by contrast, the courses are less intensive, are oriented toward practical problems, and are reinforced by practical experience. Both teachers and students spend some time working in factories or on farms,which enables them to "learn from the workers and peasants." Emphasis is placed upon "combining theory with practice" and "avoiding the three divorces," these being the divorce from practice, from the everyday concerns of workers and peasants, and from proletarian politics. Universities are guided by the "open door policy," under which workers and peasants are invited to take part in research work and intellectual institutions are concerned with the problems of production.

About one percent of middle school graduates are admitted to a university, but they must first spend at least two years in "productive labor" in a factory, commune, or army unit. Perhaps they are then considered workers, peasants or soldiers regardless of previous background, for several students so identified themselves even though their main working experience had been confined to the two years just mentioned. (Some were children of Party cadres.) Because of this confusion, we were uncertain how to interpret the claim that the numbers of workers, peasants and soldiers enrolled in the universities had increased greatly since the Cultural Revolution.

There are no formal examinations for university admission; instead, the decision is based on the applicant's "political consciousness" and "vocational competence." These are judged by interviewing the applicant and reading statements from fellow workers. At one university, the introduction of graduate work in mathematics was being considered, and we asked, "What will be the requirement for admission to graduate work?" Answer: "A knowledge of calculus." "And how will you determine whether the applicant knows calculus?" "Oh, the fellow workers will tell us!"

Though the philosophy underlying the university system was explained to us several times, we were unable to obtain a clear and complete picture of mathematics education at any one university. To some extent this was due to time limitations and to the fact that our mission was more concerned with research than with education. In any case, the table below gives some information about universities that we visited and is followed by additional details.

University	Full-time Students	Total Faculty	Math Dept Students	Math Dept Faculty	Correspondence and Short-term Students
Peking	6000	2700	500	Don't know	50,000
Tsinghua	10,000	3500*	No Mathematics Department		20,000
Futan	3000	4000+	256	111	10,000
Shanghai Normal	7000	2000	755	185	15,000
Heilungkiang	2000	1000	210	90	Don't know

*Total staff of 9000, including 3000–3500 faculty members

+Total staff of 4000, number of faculty members unknown to us

Peking University reopened in 1970–71, after the Cultural Revolution, and a small group of Computer Science students was admitted at that time. The first mathematics students were admitted in 1973 and graduated in 1976. The mathematics course, like most others, takes only three years. Thus the graduates have had 13 years of academic training in all — 5 in primary school, 5 in middle school, and 3 at the university. The Mathematics Department has three divisions: Mathematics, Computational Mathematics, and Information Theory. In the Mathematics Division there are about fifty new students per year. The first year starts with a review of middle school mathematics and then goes on to calculus and analytic geometry. Also included are physics, philosophy and English. The second year covers advanced calculus, ordinary differential equations, linear algebra, and English. At some point, mechanics and computer programming are also studied. In the third year, the fifty students split into three groups. Twenty study complex variables, calculus of variations, and electricity and magnetism; twenty study partial differential equations and the finite element method; and ten study measure theory, Fourier analysis, and eigenfunction expansions.

Tsinghua University (in Peking) is an engineering school and has no mathematics department as such. Each department teaches its own courses in calculus and differential equations, some have linear algebra, a few probability, and the electrical power engineers study complex variables. Students are taught largely through working in the university's workshops and factories under supervision of their professors. At any given time, about half the students are away from the campus, gaining practical experience in factory, commune, or army.

Futan University (in Shanghai) included Computer Science in its Mathematics Department until 1975, but now there is a separate Computer Science Department with three sections: Computational Mathematics, Information Theory, and Software. Of the 256 students in the Mathematics Department, 120 are enrolled in a mathematics program and 136 in a mechanics program. The mathematics instruction is organized around three topics: curves and surfaces in industry, the finite element method for solving partial differential equations, and mathematical methods in industrial process control. The chief orientation is toward applying mathematics and using new teaching material. For example, in 1975 some second-year students went to the oil fields to study electrical phenomena associated with exploration for oil. Instruction was organized around this project for eight months and included linear algebra, advanced calculus, the finite element method, electromagnetic field theory, and the equations of mathematical physics.

The following table, constructed for us after much discussion and consultation with students, shows a typical program for students specializing in mathematics at Futan University. In addition, there would be one month of military service, one month of harvest work, four months on a graduation project, and the vacation periods of two weeks each winter, four weeks each summer. The numbers in parentheses indicate hours of instruction per week.

2 months	Calculus (6), algorithms (4), physics (4).
3 weeks	Work in factory.
4 months	Calculus and solid geometry (6), curves and surfaces (4), physics (4).
1 month	Analyze and solve problems; self-study and small-group projects.
1 month	Work in factory and study mechanics there.
5 months	Finite element method and linear algebra (6), calculus (4), physics (2).
2 months	Solve practical problems, sometimes in factory.
4 months	Mathematical methods of industrial process control — one variable (number of hours not given). (At least 1 month would be spent in factory.)
6 months	Mathematical methods of industrial process control — several variables (number of hours not given). (At least 1 month would be spent in factory.) Probability and statistics for 6 weeks, 2 hours per week.

According to its spokesman, the Mathematics Department responded strongly to the Cultural Revolution. "Before the Cultural Revolution the Department was revisionist and its leadership not closely associated with the proletariat." But in 1968, a Mao Tse-tung propaganda team of workers marched into the Department, and under party leadership the students and teachers criticized such revisionist ideas as "knowledge is my own." In 1970, workers, peasants and soldiers were enrolled as

students. "After this revolution, the younger and middle-aged teachers made much progress and some of the older teachers also made progress. The university persists in linking education with productive labor and in carrying on open-door schooling."

Shanghai Normal University trains 3000 to 5000 middle school teachers every year and provides in-service training for both primary and middle school teachers. There are currently 7000 students living on campus. Since the Cultural Revolution, off-campus students are also enrolled — for example, 15,000 students by correspondence, students from Shanghai factories who enroll in technical training courses, and teachers from state farms and communes who enroll in teacher training courses.

In the Mathematics Department there are seven teaching groups: computational mathematics, applied mathematics, application of computers, training of middle school teachers, physics teaching, research on mathematics in middle schools, and correspondence courses. Of the 185 teachers in the department, about one fourth are women. Of the 755 students enrolled in the department, 180 take a one-year course and 575 a three-year course. Since 1974 the three-year students have been divided into three sections, specializing respectively in computational mathematics, applied mathematics, and applications of computers.

All of the regular students take courses in political theory (especially Marxism and Mao Tse-tung's teaching on education), a foreign language, Mao Tse-tung Thought, and physical education. The following additional courses are offered to the mathematics students: algebra and geometry, calculus, differential equations, algorithms, linear algebra, probability and statistics, finite element method, principles of computers, mechanical drawing, and physics. The amount of time spent on these subjects varies according to the student's specialty.

Heilungkiang University (in Harbin) has three different programs in its Mathematics Department: Mathematics, Numerical Methods, and Computer Science. The 20–30 mathematics students take algebra, calculus and analytic geometry in their first year. The second year is devoted to ordinary and partial differential equations, probability theory, and complex variables. In the middle of the second year, students split into groups of four or five for cooperative work on practical projects. A project may be based on ideas of teachers or students, a request for help from a factory, or a request from the state. One current project is in a state-run paper mill, where a computer is used to control the processing of pulp in a vat. The problem is how to control the temperature, the pressure, and the amount of alkali to maximize the production rate of paper.

The Mathematics Department faculty of 90 includes those who teach physics and English to students in the department. The high faculty-to-student ratio is explained by the fact that, in addition to teaching, the faculty engages in research (always connected with practical problems) and also in the time-consuming program of taking students to factories or the countryside for their practical work. Teachers must often go first, to explore the possibilities for projects and to prepare for a specific project. The university is reluctant to take on many students until it has codified its educational program.

Harbin Industrial College is an engineering school that has no mathematics department as such. Instead, each department has faculty members whose jobs are to teach only the basic parts of mathematics that are useful for the department's specialty. Students have no textbooks but use duplicated materials. During their three years at the Industrial College, students spend three months working in a factory, one month working on a commune, one month in the army, and three months "in society studying the political struggle."

Other forms of post-secondary mathematical education. Correspondence courses are given by some universities, and short-term and refresher courses are given by both universities and factories. This is a substantial enterprise, as can be deduced from figures in the above table and from the fact that, at almost every factory that we visited, we were told of extensive part-time classes for the workers. In general, there is great eagerness for education and the importance of mathematical training is widely recognized. Especially popular, in connection with the factories, are courses on elementary methods of operations research. See the delegation's detailed report for a description of the extraordinary Chinese effort to disseminate such methods.

Graduate work in mathematics hardly existed in China during our visit, but it was being considered at some universities. For university graduates who are selected to join the various research institutes, further training is obtained through individual study, seminars, and personal guidance from older members of the institutes.

Four questions. Our overall impressions of Chinese mathematical education are summarized in the next section. However, before describing them it should be acknowledged that we learned much from our Chinese colleagues, and our experiences in China "raised our consciousness" on a number of issues, such as the following:

(1) In our concern for the mathematically talented student, are we neglecting those students who, while not especially gifted, will be users of mathematics?

(2) Should we, in our teaching and writing, pay more attention to the relation of mathematics to other sciences, engineering, philosophy, history, economics, etc?

(3) To what extent should academic mathematicians combine theory and practice by engaging in industrial, government and military research?

(4) To what extent should motivation through applications be available in the mathematics curriculum?

Summary. Of all the facets we saw of Chinese society, education seems to have been most profoundly affected by the Cultural Revolution. The anti-elitist, anti-bourgeois campaigns produced a new educational system based on the philosophy that "education must serve proletarian politics and must be combined with productive labor." Though the current program in mathematical education was repeatedly described as experimental, that seemed to refer to details rather than general principles. Indeed, we were also told "The policy and line are not experimental any more. The general principles have been set. The problem is how to improve teaching materials." Two dominant guidelines of the new system are "combine theory with practice" and "place politics in command."

To achieve the combination of theory with practice, students and teachers engage in "productive labor" at a factory as an integral part of many mathematics courses. For example, the first-year calculus course at Peking University last year was illustrated throughout with the calculation of where to support a horizontal laser tube. Similarly, the course at Shanghai Normal was centered on the design of a cam for a weaving machine. Most of the topics taught in these courses were motivated by these applications. In each case the students and teachers discovered the problems through their experiences in factories and then, in the classroom, learned enough mathematics to provide practical solutions. Other examples of combining theory with practice were given earlier.

As has been described, placing politics in command has a profound effect on university admission policy, on the relationship of teachers to students and students to each other, and on the relationship of the academic community to the rest of society. Each week students and teachers must participate in political discussions in order to exchange ideas about the class struggle and to "grasp the correct line." In accordance with this line, the existence of genius is denied and talent is believed to have only a minor effect on a person's capacity. True dedication to serve the state is claimed to be adequate to overcome any existing differences in talent. With respect to mathematicians, it is expected that higher esteem will be accorded to those who help factory workers solve problems than to those who prove theorems in a research institute. Mathematicians are taught to think first of contributing to material progress rather than pursuing problems because they are intellectually stimulating. Current policy does not encourage the idea that a person may be inspired by the inner beauty of mathematics.

The Chinese educational system seems to have two main goals in addition to producing general basic literacy. The first and currently paramount concern is to promote egalitarianism and a social consciousness, preventing a sense of elitism among the educated. The second is to provide the training required to develop modern technology and thus turn China into an industrial state.

It is hard to judge how well the first goal is being met, though it does at least seem clear that worker-peasant origins have become something to admire rather than scorn.

Judging the effectiveness of the system in fulfilling the second goal is almost impossible. The new

courses are highly experimental and there seems to be no method to evaluate their effectiveness. Although the students are tested on their political attitudes and knowledge, it is claimed there is no comparable evaluation of their technical proficiency.

One may ask whether the present system will produce enough people with sufficient mathematical training to meet the needs of modern technology, and whether the system will be able to train university mathematics teachers and researchers. At present, most mathematicians in the People's Republic of China were trained before the Cultural Revolution, many of them abroad. How will the supply be replenished?

Acknowledgments

The first acknowledgement is to our Chinese hosts and colleagues, and to the Scientific and Technical Association of the People's Republic of China. Their excellent planning and warm hospitality enabled us to make the most of our limited time in their country.

The second acknowledgement is to the Committee on Scholarly Communications with the People's Republic of China, and to the organizations that sponsor it. Without the support at all stages of CSCPRC, and of its staff member Anne Fitzgerald, our trip would have been impossible.

The third acknowledgement is to my companions on the trip: Mac Lane, Carrier, Brown, Feit, Fitzgerald, Keller, Kohn, Leban, Pollak and Wu. All collaborated in writing the report on which this article is based. I have drawn freely upon that report, often using whole paragraphs, in other cases rearranging or rewording because of space considerations, or in order more accurately to represent my own impressions. Comments from Mac Lane and Pollak were especially helpful.

The fourth acknowledgement is to Professor Frank Swetz of the Capitol Campus of Pennsylvania State University in Harrisburg, and to his publisher, the M.I.T. Press. They have kindly permitted the use of excerpts from his book to describe the historical background of Chinese education in mathematics.

Finally, I should acknowledge that in any article of this sort, there are certain to be errors of fact and of interpretation. I shall be grateful if interested readers will tell me about them.

References

CSCPRC Report Series, No. 3, Pure and Applied Mathematics in the People's Republic of China (Anne Fitzgerald and Saunders Mac Lane, editors), National Academy of Sciences, Washington, D.C., 1977. (This is the report of our delegation. Extended excerpts from it have appeared in the present article, the article by Feit listed below, and two articles in the SIAM News: "Math in China is devoted to production and society", vol. 9, no. 6, December 1976, pp. 1, 4; "The second report on a visit to China", vol. 10, no. 2, april 1977, pp. 1, 6.)

CSCPRC Report Series, No. 1, Solid State Physics in the People's Republic of China (Anne Fitzgerald and Charles P. Slichter, editors), National Academy of Sciences, Washington, D.C., 1976.

T. E. Cheatham *et al.*, Computing in China: A Travel Report, Science, 182 (1973) 134–140.

Chandler Davis, A mathematical visit to China, Notices Amer. Math. Soc., 19 (1972) 167–169.

John de Francis, Mathematical competitions in China, this MONTHLY, 67 (1960) 756–762.

Walter Feit, A mathematical visit to China. May 1976, Notices Amer. Math. Soc., 24 (1977) 110–113.

Wolfgang K. H. Panofsky, Observations on high energy physics in China, US-China Relations Program, Stanford University, 1977.

Stephen Salaff, A biography of Hua Lo-keng, Isis, 63 (1972) 143–183.

Jack Schwartz, Computing in China: a trip report, Unpublished, 1973, 18 pages.

Frank Swetz, The Chinese Mathematical Olympiads: a case study, this MONTHLY, 79 (1972) 899–904.

———, Selected aspects of mathematics education in the People's Republic of China, Penn State University, Capitol Campus, Dept. Math. Mimeo Series, No. 5, 1972, 29 pp.

———, Mathematics Education in China: its Growth and Development, M.I.T. Press, Cambridge, 1974.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195.

THE RATIONAL CUBOID REVISITED

JOHN LEECH

1. Introduction. A *classical rational cuboid* is a rectangular parallelepiped of which the lengths of the edges and the face diagonals are all integers. If in addition the length of the body diagonal is also an integer, it is a *perfect rational cuboid.* It is a notorious unsolved problem whether any perfect rational cuboids exist; none is known, but no satisfactory proof of their impossibility has been given. If the lengths of the edges, face diagonals and body diagonal are $x_1, x_2, x_3, y_1, y_2, y_3, z$, respectively, this would require simultaneous solution in integers of the equations

$$x_2^2+x_3^2=y_1^2, \quad x_3^2+x_1^2=y_2^2, \quad x_1^2+x_2^2=y_3^2, \quad x_1^2+x_2^2+x_3^2=z^2.$$

Martin Gardner [**4**], observing that no perfect rational cuboid is known, enquired whether it is possible for any six of the seven lengths to be integers. As thus posed, this question has three variants: we may allow the body diagonal or a face diagonal or an edge to have an irrational length. Of these, the first has been discussed extensively, the second much less so, and the third scarcely at all, as far as I can trace. In the present account I deal with each in turn, giving more emphasis to the less familiar cases. I include digressions referring to the construction of spherical triangles of which the sides and angles have rational numbers for all their sines and cosines, and to the possibility of sets of four squares whose sums in pairs or whose differences are all perfect squares, before returning to the perfect cuboid.

I begin by quoting two well-known general results. In the first, let ξ, η, ζ be the sides of a rational right-angled triangle, i.e., integers satisfying $\xi^2+\eta^2=\zeta^2$. Then there exist coprime integers α, β such that

$$\xi:\eta:\zeta=2\alpha\beta:\alpha^2-\beta^2:\alpha^2+\beta^2.$$

If ξ is odd, then α, β are both odd, but if η is odd, then one of α, β is even. The effect of exchanging ξ, η is to replace α, β by their sum and difference, removing a common factor 2 when these are both even. α, β are called the *generators* of a solution of $\xi^2+\eta^2=\zeta^2$ in integers. Since we are concerned only with ratios $\xi:\eta:\zeta$, it does not matter whether ξ, η, ζ are coprime. Also, since only their squares appear, it does not matter if formulae for ξ, η give negative values, and numerical values for α, β which are negative or have $\alpha<\beta$ will be tacitly rearranged to have $\alpha>\beta>0$.

The second refers to rational points on certain plane cubic curves. Those with which we are concerned have equations which, when expressed in inhomogeneous coordinates, contain only terms of odd degree. Thus the origin is a point of inflexion of the curve and is a centre of symmetry. The rational points on the curve form an abelian group, which may be described additively. The origin is the zero element. The negative of a point is its reflection in the origin. The sum of two points is found by taking the negative of the point where their join meets the curve again. Thus the sum of any three collinear points is zero. By Mordell's famous finite basis theorem ([**13**], [**14**] ch. 16), this group of rational points is finitely generated. Thus, beginning with a finite set of rational points, we can construct all rational points on the curve by repeatedly adjoining the further intersections with the curve of the tangents at points and of the chords through pairs of points. This construction may give a finite closed set of *exceptional points*, or there may be one or more *ordinary points*, corresponding to generators of infinite order and leading to an infinity of rational points. The known sets of exceptional points are listed by Mordell ([**14**] p. 146). Among them, we note the following. When the cubic has the form

$$au(v^2-1)=bv(u^2-1),$$

there are eight exceptional points, namely the origin, three points at infinity in the directions

$u=0, v=0, u/v=a/b$ (these form a closed set of four points), and the four points $u=\pm 1, v=\pm 1$. In the special case that $a/b=\{(w^2-1)/2w\}^2$, for w rational, there are eight further exceptional points, corresponding to $u=\pm w, \pm 1/w$, each with $v=(u+1)/(u-1), (1-u)/(1+u)$. It is confidently conjectured that this latter set of sixteen exceptional points is the largest possible, and that no other set of exceptional points properly contains the foregoing set of eight points. We call this *Conjecture* C. It implies that other rational points on these curves are ordinary points.

2. Cuboid with body diagonal irrational. This is the classical rational cuboid problem; it has been widely discussed in the past (Dickson [**3**] ch. 19 refs. 1–30) and comprehensively, more recently, by Kraitchik [**6**]. We require integer solutions of the equations

$$x_2^2+x_3^2=y_1^2, x_3^2+x_1^2=y_2^2, x_1^2+x_2^2=y_3^2,$$

with no condition on $x_1^2+x_2^2+x_3^2$. Suppose that each of these equations has a solution with generators a_i, b_i, so that

$$x_{i+1}:x_{i+2}:y_i=2a_ib_i:a_i^2-b_i^2:a_i^2+b_i^2,$$

$i=1,2,3$, with subscripts reduced modulo 3. Then we seek integer solutions of

$$\frac{a_1^2-b_1^2}{2a_1b_1}\cdot\frac{a_2^2-b_2^2}{2a_2b_2}\cdot\frac{a_3^2-b_3^2}{2a_3b_3}=1. \tag{2.1}$$

When studying integer solutions, it is convenient to standardize our generator pairs to be of opposite parity, which amounts to replacing (2.1) by

$$\frac{a_1^2-b_1^2}{2a_1b_1}\cdot\frac{a_2^2-b_2^2}{2a_2b_2}=\frac{\alpha^2-\beta^2}{2\alpha\beta} \tag{2.2}$$

with no loss of generality. This has many solutions, such as

$$\frac{6^2-5^2}{2\cdot 6\cdot 5}\cdot\frac{11^2-2^2}{2\cdot 11\cdot 2}=\frac{8^2-5^2}{2\cdot 8\cdot 5}. \tag{2.3}$$

Kraitchik [**6**] gives dimensions and generators for 241 cuboids in which the edge of odd length is smaller than 10^6; in a supplement [**7**] he gives 18 cuboids (only 16 new). Lal and Blundon [**8**] list all cuboids corresponding to solutions of (2.2) with $a_1, b_1, \alpha, \beta \leqq 70$. I have deposited [**10**] a list of all solutions of (2.2) in which two of the pairs $a_1, b_1, a_2, b_2, \alpha, \beta$ do not exceed 376, giving also such of the edges of the cuboids as do not exceed 10^6. There are 62 cuboids whose body diagonals do not exceed 10^5.

Cyclic permutation of the fractions in (2.1) does not lead to any essentially different cuboids, but reversal of the cyclic order, which corresponds to exchange of the fractions on the left of (2.2), leads to a distinct, though related, cuboid. Kraitchik calls this the *derived cuboid.* For example, the numerical solution (2.3) gives the two cuboids with $x_i=240, 44, 117$ and $x_i=429, 2340, 880$, of which the former is the cuboid with the smallest dimensions. The dimensions of two mutually derived cuboids are reciprocally related, products of pairs of corresponding dimensions being equal; in this numerical example we have

$$240\cdot 429=44\cdot 2340=117\cdot 880=102960.$$

Parametric solutions have been given by various writers; several are given by Kraitchik [**6**]. The simplest, which was known to Euler, may be obtained in the following way. Any solution of (2.2) corresponds to solutions of

$$\{2ab(\alpha^2-\beta^2)\}^2+\{2\alpha\beta(a^2-b^2)\}^2=\text{sq};$$

this may be written as

$$a^2b^2\alpha^4 + a^2b^2\beta^4 + \alpha^2\beta^2a^4 + \alpha^2\beta^2b^4 - 4a^2b^2\alpha^2\beta^2 = \text{sq},$$

or as

$$a^2\alpha^2(b^2\alpha^2 + a^2\beta^2 - 4b^2\beta^2) + b^2\beta^2(b^2\alpha^2 + a^2\beta^2 - \text{sq}) = 0.$$

Any solution of $b^2\alpha^2 + a^2\beta^2 = 4b^2\beta^2$ will satisfy this, so, putting

$$b\alpha : a\beta : 2b\beta = p^2 - q^2 : 2pq : p^2 + q^2,$$

we get $\alpha : 2\beta = p^2 - q^2 : p^2 + q^2$ and $a : 2b = 2pq : p^2 + q^2$, i.e.,

$$\alpha = 2(p^2 - q^2), \quad a = 4pq, \quad \beta = b = p^2 + q^2. \tag{2.4}$$

The numerical solution (2.3) corresponds to

$$p = 2, \quad q = 1, \quad b\alpha : a\beta : 2b\beta = 3:4:5.$$

For any fixed pair of generators a, b, the equation (2.2) is equivalent to a plane cubic curve

$$\frac{a^2 - b^2}{2ab} = \frac{u^2 - 1}{2u} \cdot \frac{2v}{v^2 - 1} \tag{2.5}$$

in the inhomogeneous coordinates $u = \alpha/\beta$, $v = a_2/b_2$. This is of the special form described in §1, with eight exceptional points which correspond to trivial solutions. According to Conjecture C, non-trivial solutions always correspond to ordinary rational points on the curve; any counter-example would lead to an unprecedented set of exceptional points. Thus if a, b occur in any one solution, they occur in an infinity of solutions. But not all pairs a, b occur in solutions. It may be shown by an infinite descent argument that $a/b = 2$ is impossible; thus there is no classical rational cuboid with two edges in the ratio 3:4 of the most familiar rational right-angled triangle.

3. Cuboid with one edge irrational. Here we require two of the edges, all the face diagonals and the body diagonal to be integers. Thus we need to solve $x_1^2 + x_2^2 = y_3^2$ in combination with a positive integer t, the square of the irrational edge, such that $t + x_1^2, t + x_2^2, t + y_3^2$ are all perfect squares. This problem was posed by "Mahatma" **[12]**; readers' solutions include $x_i = 124, 957, t = 13852800$. Bromhead **[1]** extends this numerical solution to a one-parameter family of solutions.

For any integer solution of $\xi^2 + \eta^2 = \zeta^2$, the problem of finding rational numbers τ, such that $\tau + \xi^2, \tau + \eta^2, \tau + \zeta^2$ are all rational squares, is a typical example of Fermat's "triple equations." (See **[5]**, pp. 321–328, for a general discussion.) Fermat's methods lead to an infinity of rational values of τ for each set ξ, η, ζ, of which the simplest is found to be

$$\tau = (\zeta^8 - 6\xi^2\eta^2\zeta^4 + \xi^4\eta^4)/(2\xi\eta\zeta)^2. \tag{3.1}$$

(Another derivation of this solution is indicated below.) Integer solutions can be found by multiplying by the common denominator; in this example we have $x_1, x_2, y_3 = 2\xi\eta\zeta(\xi, \eta, \zeta)$ and $t = \zeta^8 - 6\xi^2\eta^2\zeta^4 + \xi^4\eta^4$. This value of t is positive only if $\zeta^2/\xi\eta > 1 + \sqrt{2}$, which requires ξ/η or η/ξ to exceed 1.8832.... The simplest real cuboid of this form is for $\xi = 5, \eta = 12$ and has $x_i = 7800, 18720, t = 211773121$.

These methods do not give all solutions for every pair ξ, η. We are looking for solutions of the equations

$$x_1^2 + x_2^2 = y_3^2, \quad z^2 = x_1^2 + y_1^2 = x_2^2 + y_2^2, \tag{3.2}$$

other than those of the trivial form with $z = y_3$, corresponding to $t = 0$. For the cuboid to be real, we require $t = z^2 - y_3^2 > 0$; let us ignore this restriction for the moment. Thus we are looking for solutions

of $z^2 = x_i^2 + y_i^2 = x_j^2 + y_j^2$ for which $x_i^2 + x_j^2$ is square, other than with $x_j = y_i$. For example, among the representations

$$65^2 = 52^2 + 39^2 = 56^2 + 33^2 = 60^2 + 25^2 = 63^2 + 16^2$$

we notice the values $60^2, 63^2$, whose sum is 87^2. This leads to the simplest solution of the equations, though unfortunately with the negative value $t = -3344$. A computer search is easily organized; we examine numbers z of the form $4n + 1$, rejecting primes and those with factors of the form $4n + 3$, and express the square of each remaining z as the sum of two squares in all possible ways. From this list we find all pairs of squares whose sum is a square distinct from z^2. I have deposited [**10**] a list constructed in this way. From this list I have found that there are 100 primitive solutions with $z < 10^5$, of which 46 have $t > 0$.

The generators for the equations (3.2) are integers $a, b, \alpha_1, \beta_1, \alpha_2, \beta_2$, such that

$$\frac{x_2}{x_1} = \frac{a^2 - b^2}{2ab}, \quad \frac{z}{x_1} = \frac{\alpha_1^2 + \beta_1^2}{2\alpha_1\beta_1}, \quad \frac{z}{x_2} = \frac{\alpha_2^2 + \beta_2^2}{2\alpha_2\beta_2},$$

which have to satisfy

$$\frac{\alpha_1^2 + \beta_1^2}{2\alpha_1\beta_1} \cdot \frac{2\alpha_2\beta_2}{\alpha_2^2 + \beta_2^2} = \frac{a^2 - b^2}{2ab}. \tag{3.3}$$

Putting $u = \alpha_1/\beta_1, v = \alpha_2/\beta_2$, we express this, for fixed x_2/x_1, as the plane cubic curve

$$x_1 v(u^2 + 1) = x_2 u(v^2 + 1), \tag{3.4}$$

which is of the general form described in §1. The trivial solution with $t = 0$ corresponds to $u = a/b, v = (a + b)/(a - b)$; as we shall see below, this is an ordinary rational point. This leads to an infinity of rational points for each ratio x_2/x_1 and so to an infinity of solutions with x_1, x_2 in this ratio. This construction reproduces the solutions obtained by Fermat's methods. The tangent at an ordinary point corresponding to $t = 0$ intersects the cubic again in a point corresponding to the solution (3.1) above.

Examination of the solutions having $z < 10^5$ shows that many of these form cycles of four. For example, among the representations of 12025^2 we find

$$12025^2 = 11655^2 + 2960^2 = 11100^2 + 4625^2 = 11440^2 + 3705^2 = 12012^2 + 559^2, \tag{3.5}$$

for which the four sums $11655^2 + 11100^2$, $11100^2 + 11440^2$, $11440^2 + 12012^2$, $12012^2 + 11655^2$ are all perfect squares. When scaled, as here, to a common value ζ, the lowest common multiple of the values of z occurring in the constituent primitive solutions, these cycles take the form of solutions of the simultaneous equations

$$\zeta^2 = \xi_i^2 + \eta_i^2, \quad i = 1,2,3,4, \tag{3.6a}$$

$$\xi_i^2 + \xi_{i+1}^2 = \text{square (including } \xi_4^2 + \xi_1^2), \tag{3.6b}$$

$$\xi_1\xi_3 = \xi_2\xi_4. \tag{3.6c}$$

This last equation (3.6c) implies that only two distinct ratios of ξ_i are involved in (3.6b), although the four ratios ζ/ξ_i in (3.6a) are all distinct. Thus each cycle includes two distinct solutions for each of the two ratios x_2/x_1 occurring in it. When these solutions are identified with points on the cubic (3.4), it is found empirically that these form pairs collinear with points identified with the trivial solution having $t = 0$.

A converse result can be proved, namely that any two rational points collinear with such a trivial point correspond to a pair of solutions forming part of a cycle of four solutions. A brute force proof may be obtained by considering a line passing through such a trivial point with an arbitrary rational

gradient. Its further intersections with the cubic are not in general rational, being instead a quadratic irrational pair. Thus although $\xi_1/\xi_2 = \xi_4/\xi_3 = x_2/x_1$ is a rational number, $\xi_2/\xi_3 = \xi_1/\xi_4$ is not in general rational. A lengthy calculation shows, however, that $\xi_2^2 + \xi_3^2$ is always a rational square, and so when the four ξ_i are all rational they form a cycle of four solutions of the form (3.6). It is also found that the solutions for the other ratio ξ_2/ξ_3 then correspond to points which are similarly related to the trivial points on the cubic curve for that ratio.

Each non-trivial solution ($t \neq 0$) belongs to two such cycles of four solutions. These may be found by joining the corresponding point on the cubic to a trivial point or to its negative (its reflection in the origin). Each of these cycles includes two solutions for each ratio b/a and two solutions for each ratio β_i/α_i. Since each non-trivial solution belongs to two cycles, these cycles enable us to construct chains of solutions having the same ratio b/a or β_i/α_i. There is only one possible end for such a chain, namely the trivial solution with $t = 0$ (seen in the cycle given below); thus the chain beginning with this solution continues indefinitely. This shows that the trivial solution corresponds to an ordinary rational point on the cubic (3.4). The points constructed in this chain are just those generated by the trivial point in the abelian group of rational points.

I do not know whether these cycles and their compounds are all the possible cycles of solutions, or whether others exist (which might include cycles of four solutions not satisfying (3.6c) or chains, formed as above, making closed loops). We shall see in §8 that perfect cuboids would include cycles of three solutions.

The simplest cycle of solutions is that including a trivial solution and the solution (3.1) above. It corresponds to the tangent at a trivial point on the cubic. Let m, n be generators for $p^2 + q^2 = r^2$; thus $p = m^2 - n^2, q = 2mn, r = m^2 + n^2$. Then the generators for $\zeta^2 = \xi_i^2 + \eta_i^2$ are

$$\frac{\beta_i}{\alpha_i} = \frac{m-n}{m+n},\quad \frac{m}{n},\quad \frac{pr}{q^2},\quad \frac{qr}{p^2},$$

and those for the two ratios ξ_i/ξ_{i+1} are

$$\frac{b}{a} = \frac{m}{n},\quad \frac{pq}{r^2}.$$

These solutions show that any rational right-angled triangle can occur either as the diagonal and the two sides of a face of a cuboid, or as the body diagonal, an edge and the opposite face diagonal; in the latter case the shorter side of the triangle can be either the edge or the face diagonal.

The following cycle corresponds to points on the cubic curve independent of the foregoing. It is a generalization of the numerical example (3.5) above, which corresponds to $m = 2, n = 1$. The generators for $\zeta^2 = \xi_i^2 + \eta_i^2$ are

$$\frac{\beta_i}{\alpha_i} = \frac{mn}{m^2 - n^2},\quad \frac{m^3 - n^3}{m^3 + n^3},\quad \frac{n(m^4 + m^2n^2 + n^4)}{m(m^4 - m^2n^2 - n^4)},\quad \frac{2m^2n^2}{m^4 - m^2n^2 - n^4},$$

and those for the ratios ξ_i/ξ_{i+1} are

$$\frac{b}{a} = \frac{mn}{m^2 + n^2},\quad \frac{n(m^4 - m^2n^2 - n^4)}{m(m^4 - m^2n^2 + n^4)}.$$

Exchange of m and n leads to a different cycle which shares a cuboid with the original cycle; these are the two cycles mentioned above to which the cuboid belongs. For $m = 2, n = 1$ the shared cuboid is that mentioned above with $t = -3344$; in this case neither cycle gives a real cuboid.

We now return to the problem of finding real cuboids with $t > 0$. Both parametric cycles given above include real cuboids for suitable values of m and n. The smallest obtainable in this way is from the first cycle, and has

$$z = 2405,\quad x_i = 1443, 1800,\quad t = 461776;$$

this is the second smallest value of z for any real cuboid. The smallest and third smallest values are for the solutions

$$z = 1105, \quad x_i = 550, 576, \quad t = 618849$$

and

$$z = 3145, \quad x_i = 969, 1480, \quad t = 6761664;$$

these belong to a cycle of four real cuboids which is not of the foregoing parametric forms, having

$$\frac{\beta_i}{\alpha_i} = \frac{1}{4}, \ \frac{9}{32}, \ \frac{19}{108}, \ \frac{3}{19} \quad \text{and} \quad \frac{b}{a} = \frac{4}{9}, \ \frac{20}{37}.$$

The solution quoted above from Mahatma's readers **[12]** is the fourth smallest solution, with $z = 3845$.

We have seen that solutions exist in which β_i/α_i or b/a has any assigned rational value (excluding the trivial values $0, \pm 1$). In either case there are an infinity of solutions, corresponding to an infinity of rational points on the appropriate cubic curve. These have limit points, and by joining these to other rational points we find that rational points are dense in the neighbourhood of any rational point. In particular, they are arbitrarily close to a rational point corresponding to $t = 0$ and on both sides of it, implying that cuboids exist for both $t < 0$ and $t > 0$. Thus real cuboids exist with any assigned ratio $(m^2 - n^2)/2mn$ for their edges or edge and opposite face diagonal.

4. Cuboid with one face diagonal irrational. Several writers have considered two related problems, some evidently without recognizing the relationship. The first is that of finding sets of three integers all pairs of which have their sums and differences squares (Dickson **[3]** ch. 15, ref. 28 and cross-references there cited). The second is that of finding three squares whose differences are squares (Dickson **[3]** ch. 19, refs. 40–45). Clearly the sums of pairs of integers in the former problem are squares satisfying the requirements of the latter, and sets of squares satisfying the latter problem (with their sides doubled if necessary) can be expressed as sums of pairs of integers satisfying the former problem.

Here we consider the problem in the cuboid form, slightly different from the foregoing, requiring solutions in integers of

$$x_1^2 + x_2^2 = y_3^2, \quad x_3^2 + x_1^2 = y_2^2, \quad x_1^2 + x_2^2 + x_3^2 = z^2, \tag{4.1}$$

with no requirement on $x_2^2 + x_3^2$. For each such solution, the integers

$$2(z^2 + x_1^2), \ 2(z^2 - x_1^2), \ 2\,|\,x_2^2 - x_3^2\,|$$

satisfy the problem as first stated; their sums and differences in pairs are the squares of $2z, 2x_1, 2y_2, 2y_3, 2x_2, 2x_3$. Similarly the squares of z, y_2, x_3 and of z, y_3, x_2 have their differences square, and thus provide a pair of solutions to the problem in its second form.

We may write the equations (4.1) in the form

$$x_2^2 + y_2^2 = z^2, \quad x_1^2 + x_3^2 = y_2^2, \quad x_1^2 + x_2^2 = y_3^2. \tag{4.2}$$

Then the generators for these equations are integers $\alpha_1, \beta_1, \alpha_2, \beta_2, \alpha_3, \beta_3$, such that

$$\frac{y_2}{x_2} = \frac{\alpha_1^2 - \beta_1^2}{2\alpha_1\beta_1}, \quad \frac{y_2}{x_1} = \frac{\alpha_2^2 + \beta_2^2}{2\alpha_2\beta_2}, \quad \frac{x_2}{x_1} = \frac{\alpha_3^2 - \beta_3^2}{2\alpha_3\beta_3},$$

which have to satisfy

$$\frac{\alpha_1^2 - \beta_1^2}{2\alpha_1\beta_1} \cdot \frac{\alpha_3^2 - \beta_3^2}{2\alpha_3\beta_3} = \frac{\alpha_2^2 + \beta_2^2}{2\alpha_2\beta_2}. \tag{4.3}$$

This condition has some resemblance to (2.2), although, with the plus sign on the right, the generator pairs α_i, β_i are not cyclically related as in (2.1). Exchange of the left factors of (4.3) gives a new

solution of (4.2), but this does not correspond to exchange of the first two equations of (4.1) so the latter exchange gives a different solution of (4.3). One's immediate reaction is that performance of these exchanges alternately would generate new solutions indefinitely. Perhaps surprisingly, it does not.

Write $u_i = \{(\alpha_i^2 - \beta_i^2)/2\alpha_i\beta_i\}^2$; then (4.3) becomes $u_1u_3 = 1 + u_2$. The recurrence relation $u_{i-1}u_{i+1} = 1 + u_i$, which was discussed by Lyness [11] in the context of finding sets of three integers whose pairs have their sums and differences square, turns out to have period 5. (See also Coxeter [2].) In terms of arbitrary non-zero values u_1, u_2, we easily find the values

$$u_3 = \frac{1+u_2}{u_1}, \quad u_4 = \frac{1+u_1+u_2}{u_1u_2}, \quad u_5 = \frac{1+u_1}{u_2}, \quad u_6 = u_1, \ldots .$$

Exchange of the first two equations of (4.1) gives a solution of (4.3) corresponding to $u_4u_2 = 1 + u_3$, and exchange of the left factors of (4.3) changes this to $u_2u_4 = 1 + u_3$, one step round from our initial solution $u_1u_3 = 1 + u_2$. Solutions thus occur in cycles of five. A list of 35 such cycles, giving all in which $\alpha_1, \beta_1, \alpha_2, \beta_2$ do not exceed 50, is given in [9]. I have deposited [10] a list of all cycles in which two of the pairs α_i, β_i do not exceed 376.

All five ratios $(\alpha_i^2 - \beta_i^2)/2\alpha_i\beta_i$, corresponding to the five u_i in Lyness's cycle, are present in any one cuboid of the present form, namely

$$\frac{y_2}{x_2} = \frac{\alpha_1^2 - \beta_1^2}{2\alpha_1\beta_1}, \quad \frac{x_3}{x_1} = \frac{\alpha_2^2 - \beta_2^2}{2\alpha_2\beta_2}, \quad \frac{x_2}{x_1} = \frac{\alpha_3^2 - \beta_3^2}{2\alpha_3\beta_3}, \quad \frac{y_3}{x_3} = \frac{\alpha_4^2 - \beta_4^2}{2\alpha_4\beta_4},$$
$$\frac{x_1z}{x_2x_3} = \frac{\alpha_5^2 - \beta_5^2}{2\alpha_5\beta_5},$$

as may be verified by applying the cyclic variants of (4.3). The last corresponds to $(x_1z)^2 + (x_2x_3)^3 = (y_2y_3)^2$, which reappears below. The five ratios may be combined to solve the five-fold composite problem of finding integers $\xi, \eta_1, \eta_2, \eta_3, \eta_4, \eta_5$ such that $\xi^2 + \eta_i^2$ and $\xi^2 + \eta_i^2 + \eta_{i+1}^2$ are all squares (including $\xi^2 + \eta_5^2 + \eta_1^2$). We have only to express the five ratios $\eta_i/\xi = (\alpha_i^2 - \beta_i^2)/2\alpha_i\beta_i$ with a common denominator ξ.

To find integer solutions of (4.1), we note that

$$x_1^2 + z^2 = y_2^2 + y_3^2 = (p^2+q^2)(r^2+s^2),$$

since an integer which is expressible as the sum of two squares in two different ways is the product of two sums of two squares. This gives

$$x_1 = ps - qr, \quad y_2 = ps + qr,$$
$$z = pr + qs, \quad y_3 = pr - qs,$$

from which we obtain

$$x_2^2 = z^2 - y_2^2 = (p^2 - q^2)(r^2 - s^2),$$
$$x_3^2 = z^2 - y_3^2 = 4pqrs.$$

Thus the products of the numerators and denominators of the ratios $(p^2 - q^2)/2pq$, $(r^2 - s^2)/2rs$ have to be perfect squares x_2^2, x_3^2. Search for suitable values of p, q, r, s affords a method of finding solutions to this problem. Euler made p, q, r, s squares individually, requiring pairs of differences of fourth powers whose product is square. Pairs of fourth powers whose sums are equal (a problem also treated by Euler) give rise to such pairs, but these are by no means the smallest. Euler listed differences of fourth powers and noticed that $3^4 - 2^4$, $9^4 - 7^4$, $11^4 - 2^4$ are all square multiples of 65, so that the product of any two is square. The first two of these pairs give a solution with $x_i = 117, 520, 756$, and generators

(4.4)
$$\frac{\beta_i}{\alpha_i}=\frac{4}{13},\ \frac{1}{13},\ \frac{1}{9},\ \frac{14}{27},\ \frac{16}{21};$$

this is the second smallest cuboid of this type. Cyclic permutation of these generator pairs gives four other cuboids, including the smallest which has $x_i = 104, 153, 672$. The whole cycle also gives the smallest solution, $\xi = 78624$, $\eta_i = 115668, 508032, 349440, 55432, 21645$, to the extended problem of making $\xi^2+\eta_i^2$ and $\xi^2+\eta_i^2+\eta_{i+1}^2$ all squares.

Another way to find integer solutions of (4.1) is to notice that

(4.5)
$$z^2 = x_2^2 + y_2^2 = x_3^2 + y_3^2,$$

so z^2 is the sum of two squares in two different ways and is thus the product of two sums of two squares. Inserting the values

$$\frac{x_3}{x_1}=\frac{\alpha_2^2-\beta_2^2}{2\alpha_2\beta_2},\quad \frac{x_2}{x_1}=\frac{\alpha_3^2-\beta_3^2}{2\alpha_3\beta_3},$$

we obtain the expression

$$z^2 = 4(\alpha_2^2\alpha_3^2+\beta_2^2\beta_3^2)(\alpha_2^2\beta_3^2+\beta_2^2\alpha_3^2).$$

Following Euler, we can find solutions of this by making each factor square separately, which means finding two rational right-angled triangles with the same area $\frac{1}{2}\alpha_2\alpha_3\beta_2\beta_3$. Diophantus (*Arithmetica*, Book V, Lemma 2 to Prop. 7; **[5]**, p. 203) solved this by showing that, if $r^2 = s^2 + st + t^2$, then we may take

$$\frac{\beta_2}{\alpha_2}=\frac{s+t}{2r},\quad \frac{\beta_3}{\alpha_3}=\frac{s}{t}.$$

To solve $r^2 = s^2 + st + t^2$ we choose l, m, n such that $l + m + n = 0$ and put $s = l^2 - m^2$, $t = m^2 - n^2$, which gives $r = \frac{1}{2}(l^2 + m^2 + n^2)$. Cyclic permutation of l, m, n is equivalent to cyclic permutation of $s, t, -(s + t)$, so these solutions come in triples. The simplest such solution has $l, m, n = 1, 2, -3$ and $r, s, t = 7, 3, 5$, giving the cycle

(4.6)
$$\frac{\beta_i}{\alpha_i}=\frac{7}{32},\ \frac{4}{7},\ \frac{3}{5},\ \frac{11}{45},\ \frac{4}{33}.$$

This cycle includes the next three smallest cuboids of this type after the two given by (4.4).

A computer search for cuboids of this type is very similar to that of §3. We seek solutions of (4.5) for which $y_2^2 - x_3^2$ is a perfect square, so again we list all representations of relevant z^2 as the sum of two squares, this time seeking squares among the differences of these squares. I have deposited **[10]** a list constructed in this way, from which I have found that there are 89 such cuboids having $z < 10^5$.

The problem of finding pairs of rational right-angled triangles of equal area is exactly that of finding pairs of ratios $(p^2 - q^2)/2pq$ whose product and quotient are squares of rational numbers, and these latter squares are indeed the squares of the ratios β_i/α_i forming adjacent pairs of generator pairs in the cycles of solution. We noticed earlier that x_2^2/x_3^2 is the product of two ratios $(p^2 - q^2)/2pq$ and $(r^2 - s^2)/2rs$. The quotient of these ratios is also, therefore, a square, say

$$\frac{p^2-q^2}{2pq}\cdot\frac{2rs}{r^2-s^2}=\frac{m^2}{n^2}.$$

Then we find that

$$2mn : m^2 - n^2 : m^2 + n^2 = x_2x_3 : x_1z : y_2y_3,$$

so m, n are the generators α_5, β_5 in the cycle of generator pairs. Thus a pair of generators α, β occurs

in cuboids of this type if and only if β^2/α^2 is expressible as the product or quotient of two ratios of the form $(p^2-q^2)/2pq$.

This affords a method of search for generators which can occur in solutions. We calculate values of $2pq(p^2-q^2)$, removing any squared factors, and pair up those with the same reduced value. Then the product and quotient of the ratios $(p^2-q^2)/2pq$ are rational squares and give pairs of generators suitable for solutions. But, although every possible pair of generators can be found in this way, it does not yield every possible solution, or even every cycle of solutions. Many cycles, such as (4.4) above, do not include pairs of generator pairs with ratios whose squares are expressible as the product and quotient of a pair of ratios of the form $(p^2-q^2)/2pq$. So this is not a comprehensive search for solutions. It can be made so if we restrict attention to pairs of ratios for which $(p^2-q^2)(r^2-s^2)$ and $4pqrs$ are squares separately, using these values to obtain x_2 and x_3 as when obtaining the cycle (4.4) above, but this is then an inefficient and laborious search method.

For a fixed pair α_2, β_2, the equation (4.3) is equivalent to a plane cubic curve

$$\frac{\alpha_2^2+\beta_2^2}{2\alpha_2\beta_2}=\frac{u^2-1}{2u}\cdot\frac{2v}{v^2-1}$$

in the inhomogeneous coordinates $u=\alpha_1/\beta_1$, $v=(\alpha_3+\beta_3)/(\alpha_3-\beta_3)$. Like (2.5), this is of the special form described in §1, with eight exceptional points which correspond to trivial solutions. According to Conjecture C, non-trivial solutions always correspond to ordinary rational points on the curve; any counter-example would lead to an unprecedented set of exceptional points. Thus if α, β occur in any one solution, they occur in an infinity of solutions.

The rational point which is found from the tangent at a given rational point is worth further attention. Any solution of (4.2) leads to a rational point on the cubic curve for α_5, β_5. The new point found from the tangent at this point gives a cycle of solutions including

$$\frac{\beta}{\alpha}=\frac{y_2-y_3}{y_2+y_3},\quad \frac{\beta_5}{\alpha_5},\quad \frac{x_3}{x_2},\quad \frac{x_1}{z};$$

the fifth pair, corresponding to

$$\{zx_1(y_2^2-y_3^2)\}^2+\{x_2x_3(y_2^2+y_3^2)\}^2=\{y_2y_3(x_2^2+x_3^2)\}^2,$$

seems to have no simple expression. This is the cycle obtained from the pair of ratios

$$(p^2-q^2)/2pq,\ (r^2-s^2)/2rs.$$

whose numerators and denominators have products x_2^2, x_3^2.

Like (2.2) but unlike (3.3), the equation (4.3) does not have non-trivial solutions for all ratios α/β. It may be shown by infinite descent arguments that there are no solutions with $\alpha/\beta=2$ or 3; thus (as with the classical cuboids of §2) no cuboid of this type has edges in the ratio 3:4. Note that there are two separate results comprised here, as there is no necessary connection between the occurrence of a pair of generators α, β in cuboids of this type and of the pair $\alpha+\beta, \alpha-\beta$ giving the same triangle the other way round. (This contrasts with (2.1), where we can replace each pair a_i, b_i by a_i+b_i, a_i-b_i to obtain the derived cuboid.)

I conclude this section by giving two further parametric expressions for ratios $(p^2-q^2)/2pq$ whose product and quotient are squares. These have

$$p=2m^2+n^2,\quad r=m^2+2n^2,\quad q=s=m^2-n^2$$

and

$$p=2m^2-n^2,\quad r=m^2-2n^2,\quad q=s=m^2+n^2$$

respectively. If, in the former, m and n are such that $m^2-n^2=3t^2$, and we replace r, s by their sum

and difference, then the values of $(p^2-q^2)(r^2-s^2)$ and $4pqrs$ are squares separately, and may be used as squares of edges x_2^2, x_3^2 of a cuboid. Each such cuboid belongs to the same cycle as one obtainable from Diophantus's rational triangles of equal area given above.

5. Digression on spherical triangles. Let a, b, c be the sides and A, B, C the angles of a spherical triangle. For the right-angled triangle with $A = \frac{1}{2}\pi$, let

$$\theta_1 = b, \quad \theta_2 = \tfrac{1}{2}\pi - C, \quad \theta_3 = \tfrac{1}{2}\pi - a, \quad \theta_4 = \tfrac{1}{2}\pi - B, \quad \theta_5 = c,$$

and for the quadrant-sided triangle with $a = \frac{1}{2}\pi$, let

$$\theta_1 = B, \quad \theta_2 = \tfrac{1}{2}\pi - c, \quad \theta_3 = A - \tfrac{1}{2}\pi, \quad \theta_4 = \tfrac{1}{2}\pi - b, \quad \theta_5 = C.$$

Then, in either case, Napier's rules take the forms

$$\sin\theta_i = \tan\theta_{i-1}\tan\theta_{i+1} \quad \text{and} \quad \sin\theta_i = \cos\theta_{i-2}\cos\theta_{i+2},$$

subscripts being taken modulo 5. Write $u_i = \cot^2\theta_i$. Then, expressed in terms of these u_i, Napier's rules become

$$1 + u_i = u_{i-1}u_{i+1} \quad \text{and} \quad u_{i-2}u_iu_{i+2} = 1 + u_{i-2} + u_{i+2}$$

respectively. The former is Lyness's cycle, and it is not difficult to verify that the latter is also of period 5 and is equivalent to the former, taken two steps at a time.

Having seen in §4 that the five u_i can be squares of rational numbers of the form $u_i = \{(\alpha_i^2 - \beta_i^2)/2\alpha_i\beta_i\}^2$, we can apply this to construct spherical triangles in the following way. Define $\theta_i = 2\arctan(\beta_i/\alpha_i)$, consistently with $\cot^2\theta_i = u_i$, and note that

$$\sin\theta_i = \frac{2\alpha_i\beta_i}{\alpha_i^2+\beta_i^2}, \quad \cos\theta_i = \frac{\alpha_i^2-\beta_i^2}{\alpha_i^2+\beta_i^2},$$

which are both rational. Then the sines and cosines of all the sides and angles of these right-angled and quadrant-sided spherical triangles are rational numbers. Explicit examples follow from the cycles of β_i/α_i given in §4.

We can also construct spherical triangles which are neither right-angled nor quadrant-sided and have rational sines and cosines of all their sides and angles. For example, from a cycle of five θ_i we can form two right-angled triangles, one having sides θ_1, θ_2 containing the right angle and one having sides θ_1, θ_5 containing the right angle. If we now abut these triangles so that the sides θ_1 coincide, and the right angles are at the same end of this side, then the compound triangle, which may be the sum or the difference of the triangles, will have rational values for the trigonometric functions of all its sides and angles. More generally, we can abut two triangles formed from different cycles of θ_i which share a common value θ; various triangles can be formed from each such pair of cycles.

This is analogous to the construction of plane Heronian triangles, whose sides and area are all integers, by abutting pairs of rational right-angled triangles. It leads to triangles of which an altitude also has its sine and cosine rational. Constructions have been given for spherical triangles whose sides and angles have rational sines and cosines but which do not necessarily have an altitude with its sine and cosine rational (Dickson [3] pp. 221–224, "rational trihedral angles").

6. Four squares whose sums in pairs are square. In §2 we considered the problem of finding three squares whose sums in pairs are square. A natural extension is to enquire whether four (or more) non-zero squares exist all of whose sums in pairs are square.

Solutions for sets of three squares may be portrayed as a graph. Each solution of (2.1) is represented by a node of order 3, the three edges joining it to nodes each of which represents a pair of generators for a rational triangle. For the present we do not distinguish between a pair of generators

a, b and the pair $a+b, a-b$ generating the same triangle the other way round. Each such node is joined to nodes of order 3 corresponding to all solutions of (2.1) in which that pair of generators occurs. According to Conjecture C (§1), a pair of generators occurring in any one solution of (2.1) occurs in infinitely many solutions. These nodes are thus of infinite order.

A set of four squares whose sums in pairs are all squares would correspond to a subgraph having the connectivity of the edges of a tetrahedron, assuming that no two pairs x_i, x_j have the ratios x_i/x_j equal. Each vertex would be a node of order 3, corresponding to a solution of (2.1). A node at the midpoint of each edge would correspond to a pair of generators common to the two solutions of (2.1) at the ends of that edge. Each face would correspond to one of the squares, each edge to a sum of two squares, and each vertex to a classical rational cuboid. Examination of lists of generators for classical cuboids has shown no example of such a subgraph, and thus no set of four squares. Indeed, it has not even exhibited any closed circuit of any length, all connected subgraphs found being trees. We may therefore speculate not merely whether any tetrahedral subgraph is possible, but whether any closed circuit is possible. While no circuit is known, there is no confusion in treating generator pairs a, b and $a+b, a-b$ as equivalent. Greater care would be needed if circuits are found.

Although no set of four squares all of whose sums in pairs are squares is known, the construction of sets of four squares with five sums of pairs square is straightforward. From any two solutions of (2.1) with a generator pair in common, we can construct cuboids with edges x_1, x_2, x_3 and x_1, x_2, x_4 with all sums of pairs of squares square except for $x_3^2+x_4^2$. Indeed, from any one solution of (2.1) we obtain a pair of mutually derived cuboids which may be combined to give the four numbers $x_1^2, x_1x_2, x_1x_3, x_2x_3$, for which the five sums $x_1^4+x_1^2x_2^2, x_1^4+x_1^2x_3^2, x_1^2x_2^2+x_1^2x_3^2, x_1^2x_2^2+x_2^2x_3^2, x_1^2x_3^2+x_2^2x_3^2$ are all squares. It seems improbable, however, that the remaining sum $x_1^4+x_2^2x_3^2$ could be square also. This would require x_1/x_2 and x_1/x_3 to be ratios of the form $(p^2-q^2)/2pq$ whose product and quotient were both of this form. This in turn would lead to a cycle of five solutions to (4.3) in which the two ratios $\alpha_1/\beta_1, \alpha_2/\beta_2$ were themselves rational numbers of the form $(p^2-q^2)/2pq$. Cycles exist in which one ratio α/β is of this form (e.g., 15/8 in Figure 1 below), but none is known in which two are of this form.

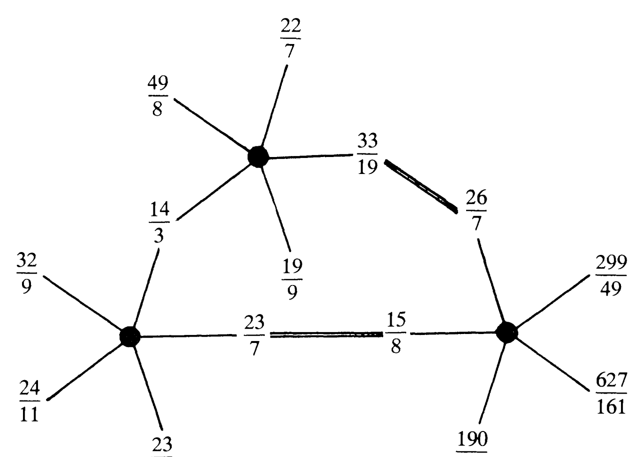


FIG. 1

7. Four squares whose differences are square. The problem of §4 being expressible as finding sets of three squares whose differences are square, we may similarly enquire whether four (or more) non-zero squares exist whose differences are all square. As in §4, we can reformulate this problem as

one of finding a sequence of four squares $x_1^2, x_2^2, x_3^2, x_4^2$, such that the sums of any two adjacent squares, of any three adjacent squares, and of all four squares, are all square. Then the sums $x_1^2, x_1^2 + x_2^2, x_1^2 + x_2^2 + x_3^2, x_1^2 + x_2^2 + x_3^2 + x_4^2$ and $x_4^2, x_3^2 + x_4^2, x_2^2 + x_3^2 + x_4^2, x_1^2 + x_2^2 + x_3^2 + x_4^2$ would form two sets of four squares whose differences were all squares.

Similarly to the graph of §6, we may portray as a graph solutions for sets of three squares whose differences are squares. This is more complicated than that of §6. First, the natural unit of the graph is the cycle of five generator pairs, which we represent by a node of order 5. Here the cyclic order is important, although the sense of rotation is not. Each solution of (4.3) corresponds to three adjacent pairs of generators, the middle pair corresponding to the right hand side of (4.3). Then we have to distinguish between the pairs of generators α, β and $\alpha + \beta, \alpha - \beta$, which we cannot regard as equivalent. As observed in §4, the occurrence of a pair α, β tells us nothing about the possibility of occurrence of the pair $\alpha + \beta, \alpha - \beta$. So, in addition to edges joining nodes for cycles to nodes for pairs of generators, we also have edges joining nodes for generator pairs to nodes for their sums and differences. According to Conjecture C (§1), each pair of generators that occurs in any cycle occurs in infinitely many cycles. The corresponding nodes are thus of infinite order. As an example, I give a subgraph in Figure 1, exhibiting the only circuit I have found by examining the cycles in the lists mentioned in §4.

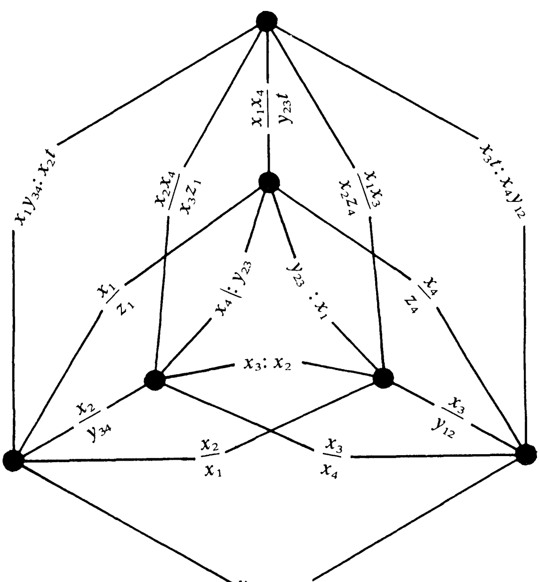


FIG. 2

If a set of four non-zero squares exists whose differences are all squares, then there would have to be six cycles of generator pairs, forming the somewhat complicated graph given in Figure 2. (This assumes that there are no extraneous equalities among the pairs of generators.) In this, instead of the generators, I give the ratios of the sides of the squares involved, as solutions of the equations

$$x_1^2 + x_2^2 = y_{12}^2, \quad x_2^2 + x_3^2 = y_{23}^2, \quad x_3^2 + x_4^2 = y_{34}^2,$$

$$x_1^2 + x_2^2 + x_3^2 = z_4^2, \quad x_2^2 + x_3^2 + x_4^2 = z_1^2,$$

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = t^2,$$

and of the consequential equations

$$x_1^2+y_{23}^2=y_{12}^2+x_3^2=z_4^2,\quad x_2^2+y_{34}^2=y_{23}^2+x_4^2=z_1^2,$$
$$x_1^2+z_1^2=y_{12}^2+y_{34}^2=z_4^2+x_4^2=t^2,$$
$$(x_2t)^2+(x_1y_{34})^2=(y_{12}z_1)^2,\quad (x_3t)^2+(x_4y_{12})^2=(y_{34}z_4)^2,$$
$$(x_1x_3)^2+(x_2z_4)^2=(y_{12}y_{23})^2,\quad (x_2x_4)^2+(x_3z_1)^2=(y_{23}y_{34})^2,$$
$$(x_1x_4)^2+(y_{23}t)^2=(z_1z_4)^2.$$

The ratio of the sides is written as a fraction, e.g. ———x_2/x_1———, where the same pair of generators belongs to both cycles (like the pair 14, 3 in Figure 1). It is written with a colon, e.g. ———$x_3:x_2$———, where a pair of generators belongs to one cycle and their sum and difference to the other (like the pairs 15, 8 and 23, 7 in Figure 1). In this latter case the ratio is to be taken as x_3/x_2 for the node to the left and as x_2/x_3 for the node to the right.

The circuit exhibited in Figure 1 can be easily seen not to form a subgraph of a graph of the form of Figure 2. Since no other circuit has been found in the graph of solutions of (4.3), this affords no indication of whether a set of four non-zero squares exists with their differences squares. On the other hand, it is possible to assign powers of 2 and of other primes to the numbers in Figure 2 consistently with the indicated relations, so the impossibility of solutions cannot be shown by simple congruence conditions. This problem thus remains completely open.

As already mentioned, there is no necessary connection between the occurrence of a pair of generators α, β in solutions of (4.3) and of the pair $\alpha+\beta, \alpha-\beta$. There are, however, many pairs α, β for which both α, β and $\alpha+b, \alpha-\beta$ occur in solutions. For example, there are an infinity of pairs α, β for which $r^2=s^2+st+t^2$ is satisfied for both $s, t=\alpha, \beta$ and $s, t=\alpha+\beta, \alpha-\beta$. The smallest of these has $\alpha, \beta=24, 11$; others are much larger.

Each pair α, β for which both α, β and $\alpha+\beta, \alpha-\beta$ occur in solutions of (4.3) leads to a sequence of four squares such that the sums of any two adjacent members and of any three adjacent members are all squares. One may speculate how long such a sequence of squares can be. For the length to exceed four, we require a sequence of cycles of solutions of (4.3) each of which includes adjacent pairs of generators α_i, β_i and $\alpha_{i+1}, \beta_{i+1}$ for which the pair $\alpha_i+\beta_i, \alpha_i-\beta_i$ belongs to one of the adjacent cycles and the pair $\alpha_{i+1}+\beta_{i+1}, \alpha_{i+1}-\beta_{i+1}$ belongs to the other, each of these latter pairs being similarly related in its own cycle. The longest such sequence I have found has the following adjacent pairs: (56,31)(17,6), (23,11)(23,7), (15,8)(26,7), (33,19)(77,19), (48,29)(35,4), (39,31)(13,9). Here (23,11) = (17 + 6, 17 − 6) etc. (Part of this sequence is seen in Figure 1.) This leads to a sequence of eight squares, with the sums of any two or three adjacent members square. There is no evident reason for the length of such a sequence to be bounded. The squares of the edges of a perfect rational cuboid would form an infinite periodic sequence. A set of four non-zero squares whose differences are all squares would lead to arbitrarily long sequences in which terms three places apart have a constant ratio; if this ratio is an integer, these sequences would be infinite.

8. The perfect rational cuboid. Having seen in §§2–4 that any six of the seven lengths involved in the cuboid can be made integers, we return to the question whether cuboids exist with all seven lengths integers. An example would satisfy simultaneously the requirements of the problems of §§2–4, and, because of the higher symmetry of the problem of the perfect cuboid, any one example would provide three examples of solutions of the problems of §§ 3,4.

One line of search for a perfect cuboid is to examine solutions of the problems involving only six of the lengths to find whether the further condition, that the seventh length be an integer, can be satisfied. For all numerical solutions so far examined, it is not satisfied, and this can be shown also for certain of the parametric solutions. An example is the solution given in §3, obtained by Fermat's method. Here we have $t=\zeta^8-6\xi^2\eta^2\zeta^4+\xi^4\eta^4$, and so $t+y_3^2=(\zeta^4-\xi^2\eta^2)^2$ and $t+2y_3^2=(\zeta^4+\xi^2\eta^2)^2$.

Hence $(\zeta^4 - \xi^2\eta^2)^4 - y_3^4 = t(\zeta^4 + \xi^2\eta^2)^2$. But the difference between two fourth powers cannot be a perfect square, as was shown by Fermat, so t cannot be square. This example is of interest in being a "near miss" to solving the perfect cuboid problem. Writing $w = \zeta^4 - \xi^2\eta^2$, we find that $x_1^2 + x_2^2$, $w^2 - x_1^2$, $w^2 - x_2^2$, $w^2 + x_1^2 + x_2^2$ are all squares, expressions differing only in the presence of the two minus signs from those for the perfect cuboid. Spohn **[15]** showed that one of the mutually derived pair of classical cuboids corresponding to (2.4) cannot be perfect; he **[16]** was unable to complete a proof of impossibility for the other. But the known parametric solutions to the partial problems are incomplete, and so results of this kind will not lead to a complete proof of impossibility; they show only that certain lines of approach cannot lead to the construction of perfect cuboids.

The existence of a perfect cuboid would imply that of a cycle of three solutions to the problem of §3, satisfying the simultaneous equations

$$\zeta^2 = \xi_i^2 + \eta_i^2, \quad i = 1,2,3,$$

$$\xi_i^2 + \xi_{i+1}^2 = \text{square (including } \xi_3^2 + \xi_1^2),$$

which may be compared with (3.6a, b), together with the condition $\xi_1^2 + \xi_2^2 + \xi_3^2 = \zeta^2$. No such cycle is known even without this added condition, the only cycles found being those of the form (3.6) and compounds of such cycles. A cycle of three solutions without the added condition would correspond to a classical rational cuboid with sides x_1, x_2, x_3 and a square z^2 such that $z^2 - x_i^2$ were all squares. This may be compared with the problem of §6, which corresponds to making $z^2 + x_i^2$ all squares. An analogous problem, soluble if perfect cuboids exist, is to find a classical rational cuboid with sides x_1, x_2, x_3 and a square z^2 such that $z^2 - x_i^2 - x_{i+1}^2$ are all squares. Search for these does not fit into the framework of the present account, and has not been made.

If a perfect cuboid exists, with dimensions as in Chapter 1, then the sets of squares $(y_k z)^2, (y_i y_k)^2, (x_i z)^2, (x_i y_k)^2$ would be such that their differences were all perfect squares, with i, j, k any permutation of 1, 2, 3. Each perfect cuboid would thus supply six such sets of four squares. These, however, are not of a completely general form; we may see this by reformulating the problem, as in §7, as one of finding sequences of four squares such that the sums of any two or three adjacent squares, or of all four squares, are all squares. In this case the four squares in the sequence are $(x_i y_k)^2, (x_i x_k)^2, (x_j x_k)^2, (x_j y_k)^2$. These have the special form that the products of the first and third and of the second and fourth are equal. Thus the problem of §7 is more general, and, should it prove soluble, this need not imply the existence of perfect cuboids.

This special form of the sequence of four squares leads to a special form of Figure 2, as each of the outer nodes of degree 5 represents the same cycle of pairs of generators as the corresponding inner node. When these nodes are identified, the resulting graph condenses to that of Figure 3 (in which the x_i etc. are as in §4, not as in §7). As with Figure 2, we see that the circuit in Figure 1 cannot form a subgraph of the form of Figure 3. Again, it is possible to assign powers of 2 and of other primes to Figure 3, showing that its possibility cannot be excluded by simple considerations of congruence.

Various relations can be inferred from the configuration of Figure 3. Thus

$$\frac{z}{x_1} \cdot \frac{x_2}{x_3} = \frac{x_2 z}{x_3 x_1} \quad \text{and} \quad \frac{z}{x_1} \cdot \frac{x_3}{x_2} = \frac{x_3 z}{x_1 x_2},$$

showing that we require a ratio of the form $(p^2 + q^2)/2pq$ and one of the form $(p^2 - q^2)/2pq$ whose product and quotient are both of the form $(p^2 - q^2)/2pq$. (Compare the requirement in §6 for two ratios of the form $(p^2 - q^2)/2pq$ whose product and quotient are of the same form. These problems are thus comparable.) This in turn would lead to a cycle of five solutions to (4.3) in which the ratios α_1/β_1 and α_2/β_2 were of the respective forms $(p^2 + q^2)/2pq$ and $(p^2 - q^2)/2pq$. Ratios α/β of these forms occur separately in cycles (e.g., 5/3 in (4.6) and 15/8 in Figure 1), but no cycle is known which combines such ratios. Any such cycle would lead to a pair of rational triangles $\xi_i^2 + \eta_i^2 = \zeta_i^2$ such that $(\xi_1^2 \xi_2 \eta_2)^2 + (\eta_1 \zeta_1 \zeta_2^2)^2$ is also square.

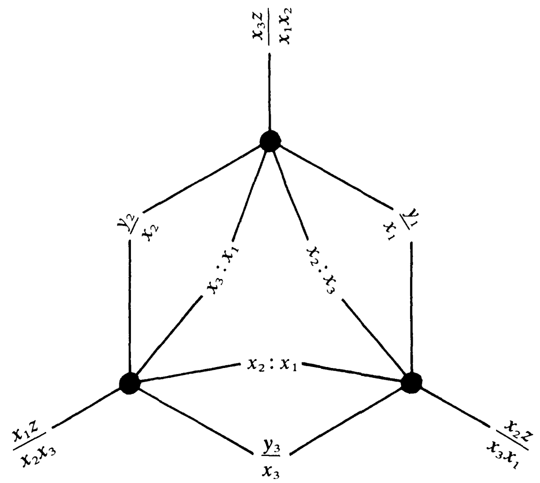


FIG. 3

We see from (2.1) and Figure 3 that, for each ratio $x_i/x_j = (a^2-b^2)/2ab$ in a perfect cuboid, the ratio of any two of $a^2-b^2, 2ab, a^2+b^2$ must be expressible as the product of two ratios of the form $(p^2-q^2)/2pq$. It appears empirically, from lists [**10**] of solutions of (2.1) and (4.3), that a rather small proportion of values of a/b can occur in solutions of these equations, unlike (3.3) in which any ratio of a/b is possible. (This is not claimed as proved. The only proved results known to me show the impossibility of $a/b = 2$ or 3 as mentioned in §§2, 4.) An even smaller proportion of values a/b are found for which all three ratios of $a^2-b^2, 2ab, a^2+b^2$ are expressible as products of ratios $(p^2-q^2)/2pq$, of which $a/b = 13/4$ is the simplest found; none has been found that leads to a perfect cuboid. A perfect cuboid would correspond to a solution of (2.1) in which all three ratios a_i/b_i have this property. No two ratios with this property have as yet been identified in any one solution of (2.1).

We saw in §4 that pairs x_i^2, x_j^2 are products of the numerators and denominators of ratios of the form $(p^2-q^2)/2pq$. In a perfect cuboid, x_i/x_j would also be of this form, which requires simultaneous solution of

(8.1a) $$(p^2-q^2)(r^2-s^2) = x_i^2 = h^2(a^2-b^2)^2,$$

(8.1b) $$4pqrs = x_j^2 = h^2(2ab)^2.$$

The factor h^2 can be absorbed by replacing p, q, a, b by hp, hq, ha, hb. We can further multiply by factors which will make $pq = rs = ab$; the problem is then one of finding an integer N which admits three factorings $N = p_iq_i$ such that the differences of squares $p_i^2 - q_i^2$ are in geometric progression. This in turn admits reformulation as requiring non-trivial integer solutions of

$$(a^2c^2-b^2d^2)(a^2d^2-b^2c^2) = (a^2b^2-c^2d^2)^2.$$

Any such solution would lead to the construction of a perfect cuboid.

Writing $u = (p+q)/(p-q)$, $v = r/s$, we see that the equations (8.1) imply

$$(2ab)^2u(v^2-1) = (a^2-b^2)^2v(u^2-1),$$

which is of the special form, with sixteen exceptional points, given in §1. However, although non-trivial rational points occur on this curve for suitable values of a, b (for example if $a/b = 4$ we have $u = 15/13$, $v = 13/8$ and $u = 25/19$, $v = 19/8$), these do not in general correspond to solutions of (8.1).

9. Problems outstanding. Perhaps the principal outstanding problem, the main theme of this account, is whether perfect rational cuboids exist. Equivalent problems, which might be more amenable to solution or proof of impossibility, include the following.

(9.1) Do cycles of solutions of (4.3) exist which are related as in Figure 3?

(9.2) Is there a ratio of the form $(p^2+q^2)/2pq$ and one of the form $(p^2-q^2)/2pq$ whose product and quotient are both of the form $(p^2-q^2)/2pq$?

(9.3) Are there non-trivial integer solutions of $(a^2c^2-b^2d^2)(a^2d^2-b^2c^2)=(a^2b^2-c^2d^2)^2$?

Related problems, which are soluble if perfect cuboids exist, include the following.

(9.4) Is there a classical rational cuboid with sides x_1, x_2, x_3 and a square z^2 such that $z^2-x_i^2$ are all squares?

(9.5) Are there sets of four non-zero squares whose differences are all squares?

(9.6) Do cycles of solutions of (4.3) exist which are related as in Figure 2?

(9.7) Is there a cycle of solutions of (4.3) in which $\alpha_1/\beta_1=(p^2-q^2)/2pq$, $\alpha_2/\beta_2=(r^2+s^2)/2rs$?

(9.8) Are there infinite or arbitrarily long sequences of squares such that all sums of two or three adjacent members are square?

Other problems include the following.

(9.9) Are there two ratios of the form $(p^2-q^2)/2pq$ whose product and quotient are both of this form?

(9.10) Are there cycles of solutions of (4.3) in which two of the ratios α_i/β_i are of the form $(p^2-q^2)/2pq$?

(9.11) Are there sets of four non-zero squares whose sums in pairs are square?

(9.12) What circuits exist in the graphs of solutions of (2.1) and (4.3)?

(9.13) Are there cycles of solutions of (3.3) other than those of the form (3.6) and compounds of these?

(9.14) What ratios other than 3 : 4 cannot occur as ratios of edges of cuboids of the types of §§2, 4?

(9.15) Is Conjecture C (§1) valid?

I hope that the present account will stimulate thought on these problems.

References

1. T. Bromhead, On square sums of squares, Math. Gazette, 44 (1960) 219–220.

2. H. S. M. Coxeter, Frieze patterns, Acta Arith., 18 (1971) 297–310.

3. L. E. Dickson, History of the Theory of Numbers, vol. 2, Diophantine Analysis, Carnegie Institute, Washington, 1920. Reprinted: Stechert, New York, 1934; Chelsea, New York, 1952, 1966.

4. M. Gardner, Mathematical games, Scientific American no. 1, 223 (July 1970) p. 118. Correction, *ibid.* no. 3 (September 1970) p. 218.

5. T. L. Heath, Diophantus of Alexandria, Cambridge, 1910.

6. M. Kraitchik, Théorie des nombres, t. 3, Analyse Diophantine et applications aux cuboides rationnels, Gauthier-Villars, Paris 1947.

7. ———, Sur les cuboides rationnels, Proc. Int. Congr. Math., Amsterdam, 2 (1954) 33–34.

8. M. Lal and W. J. Blundon, Solutions of the Diophantine equations $x^2+y^2=l^2$, $y^2+z^2=m^2$, $z^2+x^2=n^2$, Math. Comp., 20 (1966) 144–147.

9. J. Leech, The location of four squares in an arithmetic progression, with some applications. Computers and Number Theory, ed. A.O.L. Atkin and B. J. Birch, Academic Press, London and New York, 1971, pp. 83–98.

10. ———, Tables relating to rational cuboids, Math. Comp. (to appear).

11. R. C. Lyness, Cycles, Math. Gazette, 26 (1942) 62, *ibid.* 29 (1945) 231–233, *ibid.* 45 (1961) 207–209.

12. "Mahatma", Mahatma's problems no. 78, The AMA [J. of the Assistant Masters' Association, London], 44 (1949) 188. Solutions, *ibid.* 225.

13. L. J. Mordell, On the rational solutions of the indeterminate equations of the third and fourth degrees, Proc. Camb. Phil. Soc., 21 (1922) 179–192.

14. ———, Diophantine Equations, Academic Press, London and New York, 1969.

15. W. G. Spohn, On the integral cuboid, this MONTHLY, 79 (1972) 57–59.

16. ———, On the derived cuboid, Canad. Math. Bull., 17 (1974) 575–577.

DEPARTMENT OF COMPUTING SCIENCE, UNIVERSITY OF STIRLING, STIRLING, SCOTLAND.

LARVAL FISH, POWER PLANTS, AND BUFFON'S NEEDLE PROBLEM

SHERWOOD F. EBEY AND JOHN J. BEAUCHAMP

1. Introduction. Buffon's needle problem is a classical problem included in many introductory courses in probability. In addition to the traditional formulation of the problem where a needle of unit length is thrown at random onto a plane partitioned into strips of parallel unit width (see paragraph II.8 of [1]), there have been many modifications and extensions of the problem (e.g., [3], [4], pp. 251–257 of [5], pp. 37–44 of [2]). However, the authors were recently presented with a problem that seems to be a variation of the Buffon needle problem not yet considered. The general context of the problem, which resulted from consultations with members of the Environmental Sciences Division at Oak Ridge National Laboratory, may be described as follows.

At electrical power producing plants, both nuclear and non-nuclear, large volumes of water must be used for cooling in the condensation portion of the power production cycle; in addition, in nuclear plants, water is used to cool the reactor. The water is usually taken from an adjacent body of water into the plant by means of large intake pumps. These intake pumps have large impeller blades that determine the rate of flow and volume of water being used. As the water is taken into the plant, it may contain small larval fish that have passed through the preliminary screening devices to keep fish from being taken into the plant. Since the larval fish have been taken into the intake pump, they may be killed by the impeller blades if they are hit by any one of the blades. The purpose of this paper is to show how the classical Buffon needle problem may be extended to answer the following question:

"What is the probability of a larval fish being killed by an impeller blade of a pump in a power plant intake?"

2. Assumptions. The following assumptions are based upon the description of the physical problem that had been given to us and are necessary to the development of a model:

Assumption 1: The velocity of the water is such that the fish is floating with the stream and not swimming. There is turbulence in the water so that the spatial orientation of the fish may be assumed to be completely random.

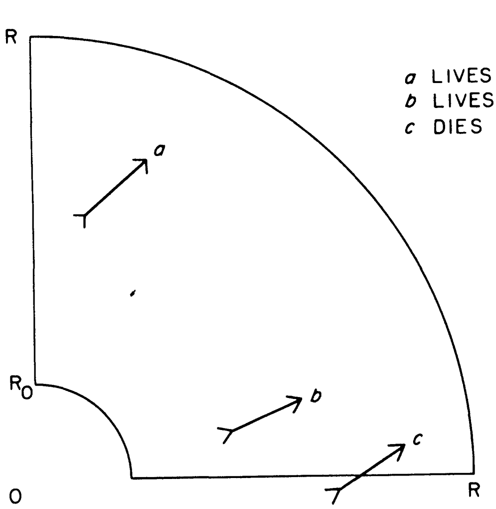


FIG. 1. Minimum Killing Region

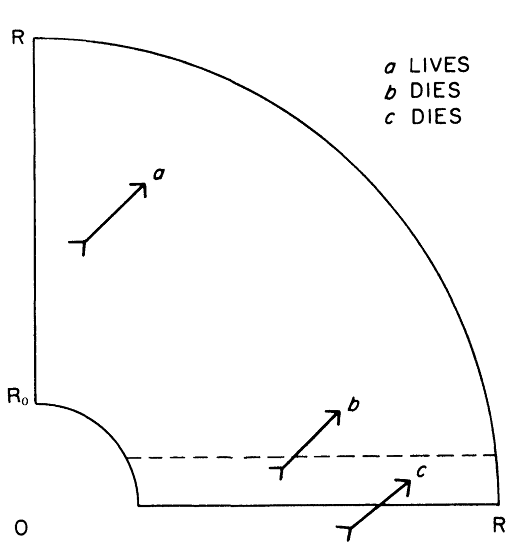


FIG. 2. Maximum Killing Region

Assumption 2: The density of the fish equals the density of the water, so it is reasonable to assume the fish are uniformly distributed through the water.

Assumption 3: It is reasonable to think of each fish as a line segment.

Assumption 4: There is uncertainty about the size of the killing region. The first possibility is that the killing region consists of line segments that are forward projections of the cutting edges of the blades. (See Figure 1.) A second possibility is that the killing region consists of bands that are the width of a vane. (See Figure 2.) The first possibility should provide a minimum probability of death due to the pump and the second should give a maximum probability.

In Figures 1 and 2, R_0 and R are the radii of the hub and intake tube, respectively. These are diagrams of a quadrant of a cross-section of the intake tube, assuming 4 blades.

Assumption 5: The probability of a fish being at a particular point of the cross-section of the pump intake tube is proportional to the velocity of the stream at that point. However, the velocity may be assumed to be constant. This was considered to be a reasonable assumption by the proposers of the problem.

Assumption 6: The impeller has a hub whose radius is not negligible.

3. Some preliminaries. In this paper we are considering the following two problems:

1. Probability of death when the killing region is a line (see Figure 1) and
2. Probability of death when the killing region is a band the width of the vanes (see Figure 2).

We shall be discussing the former case first.

The solution of this probability problem involves three steps:

1. construction of a mathematical model of the problem
2. representation of the required probabilities as integrals, and
3. evaluation of the integrals.

For our discussion, in addition to R and R_0 defined above, we let L be the length of larval fish and n be the number of impeller blades.

At first we shall suppose that $n = 4$ and at the conclusion of the problem we will adjust the answer for the case of an arbitrary number of blades.

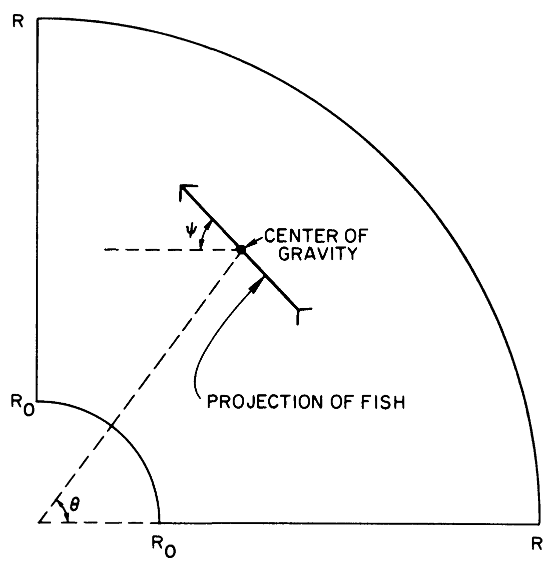


FIG. 3. Representation of variables (θ, ψ)

4. The probability model. In order to describe a "random" fish coming through the tube, we shall locate its center of gravity in the cross-section of the tube and then describe its spatial orientation. By symmetry we can suppose the center of gravity of the fish is in quadrant I. We will then locate the fish's center of gravity by polar coordinates (r, θ), where $R_0 \leqq r \leqq R$ and $0 \leqq \theta \leqq \pi/2$. Then we will observe an angle ψ made by the projection of the fish in the cross-section, where ψ will be the acute angle the fish's projection makes with the horizontal direction of the tube's cross-section. (See Figure 3.) It follows that $0 \leqq \psi \leqq \pi/2$.

Finally, we shall observe the length l of the projection of the fish on the tube's cross-section. Let ϕ = acute angle of the fish with the direction of the stream. Then $l = L \sin \phi$ and $0 \leqq l \leqq L$. (See Figure 4.)

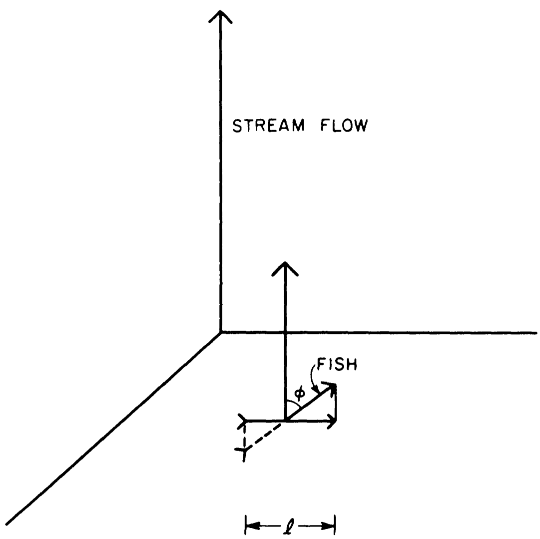


FIG. 4. Representation of variables (l, ϕ)

In summary, there are four random variables involved in obtaining the probability that a fish is killed. These are r, θ, ψ, and l. The next stage in constructing a mathematical model is to write down the probability density functions for each of these four random variables, which will be denoted f_r, f_θ, f_ψ, and f_l, respectively.

The following assumptions are consequences of Assumptions 1, 2, and 5 that have been mentioned previously and are needed for the derivation of f_r, f_θ, f_ψ, and f_l:

Assumption 7: It is reasonable to assume that the distributions of θ and ψ are both uniform, i.e., θ and ψ are each chosen at random from the interval $[0, \pi/2]$. Therefore, $f_\theta = f_\psi = 2/\pi$.

Assumption 8: It is reasonable to assume that the distribution of the angle ϕ mentioned above is uniform. Since $l = L \sin \phi$, it can be proved that the probability density function for l must be $f_l = 2/(\pi\sqrt{L^2 - l^2})$. (See Section I.10 of [**1**].)

Assumption 9: Since the location of the fish's center of gravity is randomly located in the region $R_0 \leqq r \leqq R$, $0 \leqq \theta \leqq \pi/2$, the density function for the random variable r must be Kr where K is chosen so that

$$\int_{R_0}^{R} Kr\,dr = 1.$$

Hence, $f_r = 2r/(R^2 - R_0^2)$.

In order to obtain the joint probability density function for the four variables r, θ, ψ, and l, it is necessary to make a final assumption:

Assumption 10: The four variables, r, θ, ψ, and l, operate independently so that f, the joint probability density function of the four variables, will be the product of the four density functions above. Hence,

$$f(r,\theta,\psi,l)=\frac{16}{\pi^3(R^2-R_0^2)}\cdot\frac{r}{\sqrt{L^2-l^2}} \tag{1}$$

is the required density function where $R_0 \leqq r \leqq R, 0 \leqq \theta \leqq \pi/2, 0 \leqq \psi \leqq \pi/2$, and $0 \leqq l \leqq L$. Formally, the answer to the probability problem posed is given by

$$\iiiint\limits_{\text{The Killing Region}} f(r,\theta,\psi,l)\,dr\,d\theta\,d\psi\,dl. \tag{2}$$

The next step toward a solution must be a mathematical description of the killing region.

5. The minimum killing region. For a fish whose center of gravity is in quadrant I, there are two killing regions, one represented by the horizontal axis and the other by the vertical axis. We will compute the probability for the former and then multiply by 2. (We are here using a geometric symmetry that is present in the problem.) We assume the fish will be killed if any portion of its projection on the cross-section plane crosses the horizontal axis, which represents the forward projection of the cutting edge of an impeller blade.

If r, θ, ψ, l have the meanings given above, then the y-component of point T, the projection of the fish's tail (or head), is $r \sin\theta - (l/2)\sin\psi$. (See Figure 5.)

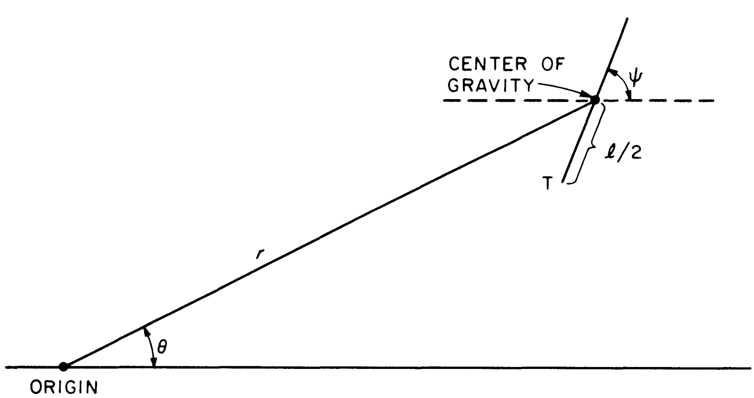


FIG. 5. Representation of variables (r, θ, ψ, l) in tube cross-section

Hence, the horizontal killing region is defined by the inequality

$$r\sin\theta-(l/2)\sin\psi<0 \quad \text{or} \quad r\sin\theta<(l/2)\sin\psi.$$

In order to help us picture this killing region in 4-dimensional space, assume that l, ψ correspond to polar coordinates for a u, v plane. Then we have the sample space (all possibilities for r, θ, l, ψ) as the cross-product of the two plane regions shown in Figure 6.

The relation $r \sin\theta < (l/2)\sin\psi$ is equivalent to $y < \frac{1}{2}v$ where $y = r \sin\theta$ and $v = l \sin\psi$.

The killing region (horizontal) is defined as follows: after l, ψ are arbitrarily chosen, then $v = l \sin\psi$ is known and y must be less than $\frac{1}{2}v$.

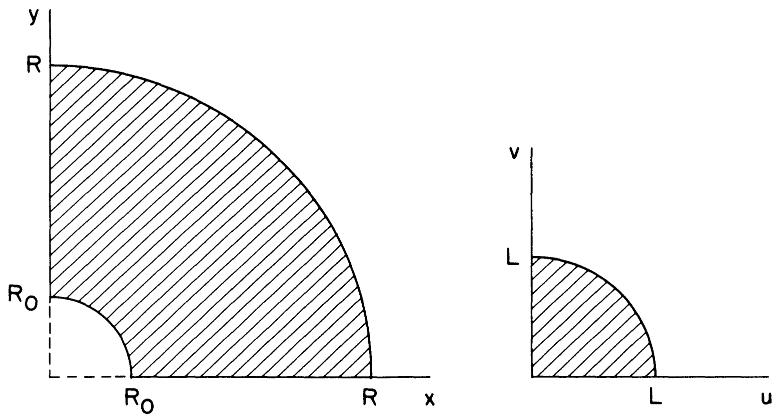


FIG. 6. Factored picture of sample space

Inequalities that describe the killing region are given by:

$$(3)\qquad \left\{\begin{array}{c} 0 \leqq l \leqq L \\ 0 \leqq \psi \leqq \pi/2 \\ 0 \leqq \theta \leqq K_1 \\ R_0 \leqq r \leqq R \end{array}\right\} \quad \text{or} \quad \left\{\begin{array}{c} 0 \leqq l \leqq L \\ 0 \leqq \psi \leqq \pi/2 \\ K_1 \leqq \theta \leqq K_2 \\ R_0 \leqq r \leqq K_3 \end{array}\right\}$$

(Designated by region I in Figure 7) (Designated by region II in Figure 7)

where the expressions K_1, K_2, K_3 are defined as follows:

$$K_1 = \operatorname{Sin}^{-1}\left(\frac{l \sin \psi}{2R}\right), \quad K_2 = \operatorname{Sin}^{-1}\left(\frac{l \sin \psi}{2R_0}\right), \quad K_3 = (l \sin \psi)/2 \sin \theta.$$

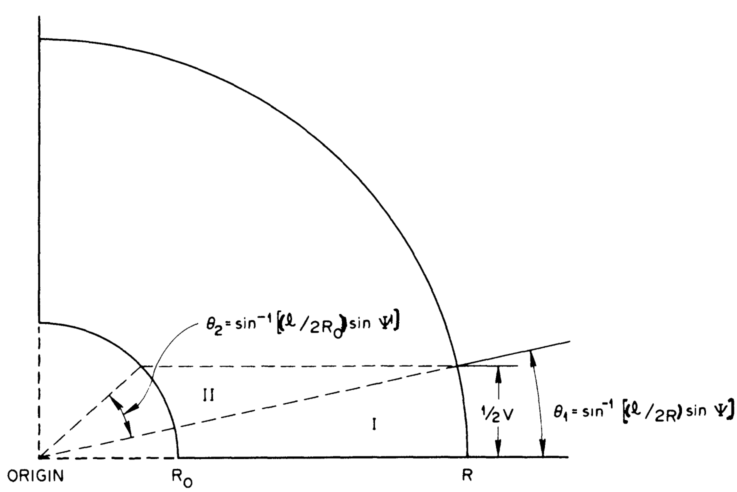


FIG. 7. Determination of horizontal "Killing region"

Consequently, we have that $\frac{1}{2}\cdot$Probability a fish is killed = Integral I + Integral II, where

$$\text{Integral I} = \frac{16}{\pi^3(R^2 - R_0^2)}\int_0^L \frac{1}{\sqrt{L^2 - l^2}}\int_0^{\pi/2}\int_0^{K_1}\int_{R_0}^{R} r\,dr\,d\theta\,d\psi\,dl \tag{4}$$

and

$$\text{Integral II} = \frac{16}{\pi^3(R^2 - R_0^2)}\int_0^L \frac{1}{\sqrt{L^2 - l^2}}\int_0^{\pi/2}\int_{K_1}^{K_2}\int_{R_0}^{K_3} r\,dr\,d\theta\,d\psi\,dl. \tag{5}$$

In order to complete the evaluation of integrals I and II, we used some approximations which can be justified by consideration of the values of the parameters L, R, and R_0 in the physical situation we were modeling. These details are described in the appendix to the paper. Then, after the evaluation of integrals I and II, doubling their sum to account for the horizontal and vertical killing regions, and extensive algebraic simplification, the (approximate) probability for death is given by the formula:

$$P = \frac{8}{\pi^3}\left(\frac{L}{R}\right)\left[\frac{R^2}{(R^2 - R_0^2)}\right]\left[2 - \overset{*}{\frac{1}{12}\left(\frac{L}{R}\right)^2} - \left(\frac{R_0}{R}\right)\left[2 + \overset{*}{\frac{5}{54}\left(\frac{L}{R_0}\right)^2} - \overset{*}{\frac{1}{180}\left(\frac{L}{R_0}\right)^4}\right]\right]. \tag{6}$$

For a further simplification, we may omit the terms marked with asterisks (*) since the magnitude of these terms will be significantly smaller than the magnitude of the remaining terms. The formula simplifies to

$$P = 16L/\pi^3(R + R_0), \quad \text{(4 blades)}. \tag{7}$$

Next, we need to consider the effect of changing the number of blades from 4 to n. The killing region will not be affected. The only change will be in the probability density function for θ, which must be changed from the reciprocal of $2\pi/4 = \pi/2$ to the reciprocal of $2\pi/n$. The formula for the probability (approximate) of a fish to die on the cutting edge of the impeller blade is then given by

$$P = 4\,nL/\pi^3(R + R_0). \tag{8}$$

6. The maximum killing region. We shall now discuss the second problem: the probability of a fish being killed by the pump if the killing region is a band that is the width of a vane of the impeller. We introduce a new parameter D, which is the width of a vane. (See Figure 8.)

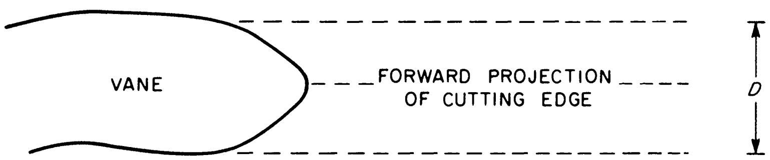


FIG. 8. Representation of parameter D in maximum "Killing region"

The meaning of variables r, θ, ψ, and l will be as before and the probability density function defined by (1) will be the same. (We are again using $n = 4$.) However, now the killing region will be defined by the inequality

$$r\sin\theta - (l/2)\sin\psi < D/2 \tag{9}$$

and the limits on l, ψ, θ, and r are as follows:

$$0 \leqq l \leqq L,\ 0 \leqq \psi \leqq \pi/2,\ 0 \leqq \theta \leqq K_4,\ R_0 \leqq r \leqq R,$$

or

$$0 \leqq l \leqq L,\ 0 \leqq \psi \leqq \pi/2,\ K_4 \leqq \theta \leqq K_5,\ R_0 \leqq r \leqq K_6,$$

where the expressions K_4, K_5, K_6 are defined as follows:

$$K_4 = \mathrm{Sin}^{-1}((D + l \sin \psi)/2R),$$
$$K_5 = \mathrm{Sin}^{-1}((D + l \sin \psi)/2R_0),$$
$$K_6 = (D + l \sin \psi)/2 \sin \theta.$$

If we let P' denote the probability of a fish being killed in this situation, then

$$\tfrac{1}{2} \cdot P' = \text{Integral III} + \text{Integral IV, where}$$

(11) $$\text{Integral III} = \frac{16}{\pi^3(R^2 - R_0^2)} \int_0^L \frac{1}{\sqrt{L^2 - l^2}} \int_0^{\pi/2} \int_0^{K_4} \int_{R_0}^{R} r\,dr\,d\theta\,d\psi\,dl$$

and

(12) $$\text{Integral IV} = \frac{16}{\pi^3(R^2 - R_0^2)} \int_0^L \frac{1}{\sqrt{L^2 - l^2}} \int_0^{\pi/2} \int_{K_4}^{K_5} \int_{R_0}^{K_6} r\,dr\,d\theta\,d\psi\,dl.$$

The remainder of the problem is to evaluate these two integrals.

Again, in order to complete the evaluation of these integrals we resorted to approximations, the details of which are in the appendix. After completing the evaluations, doubling the sum to account for both killing regions in the first quadrant, and algebraically simplifying, the answer turns out to be

(13) $$\begin{aligned} P' = {} & \frac{16}{\pi^3}\left[\frac{L}{2R} + \frac{D\pi^2}{8R}\right] + \frac{16R^2}{(R^2 - R_0^2)\pi^3}\left[\frac{L}{2R} + \frac{D\pi^2}{8R}\right] + \frac{16R_0^2}{(R^2 - R_0^2)\pi^3}\left[\frac{L}{2R} + \frac{D\pi^2}{8R}\right] \\ & - \frac{1}{R_0(R^2 - R_0^2)\pi^3}\left[(8R_0^2 D - \overset{*}{D^3})\frac{\pi^2}{4} + (8R_0^2 - 3\overset{*}{D^2})L - \frac{3}{16}\overset{*}{L^2}\pi^2 - \frac{8}{9}\overset{*}{L^3}\right] \\ & - \frac{1}{3R_0(R^2 - R_0^2)\pi^3}\left[(24R_0^2 D + \overset{*}{D^3})\frac{\pi^2}{4} + (24R_0^2 + 3\overset{*}{D^2})L + \frac{3}{16}D\overset{*}{L^2}\pi^2 + \frac{8}{9}\overset{*}{L^3}\right]. \end{aligned}$$

For a further simplification we will choose to ignore the terms that have been marked with an asterisk (*) since the magnitude of these terms will be relatively small and they will contribute very little to the final probability. When these smaller terms are ignored, the formula becomes

(14) $$P' = 4(4L + D\pi^2)/\pi^3(R + R_0).$$

Finally, if we replace 4 blades by n blades, the probability (approximate) becomes

(15) $$P' = n(4L + D\pi^2)/\pi^3(R + R_0).$$

Appendix. The following simplifications were used in the derivation of equation (6):

a. At a certain place we encountered

$$\int_0^{\pi/2} \mathrm{Sin}^{-1}\{(l \sin \psi)/2R\}\, d\psi.$$

Now $(l \sin \psi)/2R \leqq L/2R$ and $L/2R$ is very small (say < 0.10). This upper bound for $L/2R$ was suggested by estimates given to us by the members of the Environmental Sciences Division of ORNL who proposed the problem to us. Because of this we replaced $\mathrm{Sin}^{-1}\{(l \sin \psi)/2R\}$ by $(l \sin \psi)/2R$ since $\mathrm{Sin}^{-1}(t)$ is approximately equal to t when t is near 0.

b. At another place in the integration we encountered

$$\int_0^L \frac{4R^2-l^2}{l} \ln\left(\frac{1+(l/2R)}{1-(l/2R)}\right) dl.$$

Again, since $l/2R < L/2R < .10$, we replaced

$$\ln\left(\frac{1+(l/2R)}{1-(l/2R)}\right)$$

with $2(l/2R)$ since $\ln((1+t)/(1-t))$ is approximately equal to $2t$ when t is near 0.

c. We also encountered the integrals

$$\int_0^{\pi/2} \operatorname{Sin}^{-1}\left(\frac{l}{2R_0}\sin\psi\right) d\psi \quad \text{and} \quad \int_0^L \frac{4R_0^2-l^2}{l} \ln\left(\frac{1+(l/2R_0)}{1-(l/2R_0)}\right) dl.$$

From the prospective values of L and R_0, we have $(l\sin\psi)/2R_0 < L/2R_0 < .25$. However, $L/2R_0$ is not as close to 0 as $L/2R$. Hence, in evaluating these integrals, we used an additional term from the Taylor's Series expansion for $\operatorname{Sin}^{-1}(t)$ and $\ln((1+t)/(1-t))$. Specifically, we used $\operatorname{Sin}^{-1}(t) = t + (1/6)t^3$ and $\ln((1+t)/(1-t)) = 2t + (2/3)t^3$. In each case the next term of the Taylor's Series would involve t^5 and would be negligible for values of t of the magnitude $L/2R_0$.

The derived formula for P, (8), should be trustworthy as long as $L/2R$ is very small (say $< .10$) and $L/2R_0$ is small (say $< .25$).

In evaluating integrals III and IV we encountered integrals in which we had to deal with $\operatorname{Sin}^{-1}(t)$ or $t\sqrt{1-t^2}$ where t was either $(D + l\sin\psi)/2R$ or $(D + l\sin\psi)/2R_0$. In the former case we assumed that $(D + l\sin\psi)/2R$ is very small so that a linear approximation could be substituted for the function. In the latter case, we assumed that $(D + l\sin\psi)/2R_0$ was small, but of such a magnitude that including an additional term (of degree 3) in the approximation was necessary. Hence, the same cautions mentioned earlier should apply here.

Acknowledgements. This work was done while the first author was on sabbatical leave from The University of the South and a visitor at ORNL. This research was sponsored by the U. S. Energy Research and Development Administration under contract with Union Carbide Corporation, Nuclear Division.

References

1. W. Feller, An Introduction to Probability Theory and Its Applications, Vol. II, 2nd ed., Wiley, New York, 1971.
2. B. V. Gnedenko, The Theory of Probability, Chelsea, New York, 1963.
3. J. F. Ramaley, Buffon's needle problem, this MONTHLY, 76 (1969) 916–918.
4. E. F. Schuster, Buffon's needle experiment, this MONTHLY, 81 (1974) 26–29.
5. J. V. Uspensky, Introduction to Mathematical Probability, McGraw-Hill, New York, 1937.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF THE SOUTH, SEWANEE, TN 37375.
UNION CARBIDE CORP., NUCLEAR DIVISION, P. O. BOX Y, BUILDING 9704–1, OAK RIDGE, TN 37830.

ADDENDUM TO "APPORTIONMENT SCHEMES AND THE QUOTA METHOD"

(This MONTHLY, 84 (1977) 450–455)

M. L. BALINSKI AND H. P. YOUNG

The galley proofs to this article were received late and certain changes and corrections sent by the authors to the Editorial office of the MONTHLY were not included in the published report.*

Formula (1) should read

$$\max_i \frac{p_i}{a_i + 1/2} \leqq \min_i \frac{p_i}{a_i - 1/2} \text{ for } a_i \geqq 1.$$

Garrett Birkhoff communicated various suggestions which clarified our interpretation of his results. It is, regrettably, too late to incorporate his suggestions. However, it is important to note that by "entitled to 1.1. representatives" (see *SD* discussion in Section 6 of our paper) he meant $p_i/\sigma \geqq 1.1$. He also informed us that, in his discussion of "bias," the set of populous and nonpopulous states for comparing **Q** and **H** were chosen as those "...for which **Q** deviated from most of the other standard Apportionment Schemes" quoted in our papers and were not "conveniently chosen in each case."

GRADUATE CENTER, THE CITY UNIVERSITY OF NEW YORK, 33 WEST 42 STREET, NEW YORK, NY 10036, AND THE INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS, LAXENBURG, AUSTRIA

* The June-July 1977 issue of the MONTHLY was already printed when the Editorial office received the authors' changes and corrections.—*Editor.*

MATHEMATICAL NOTES

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Department of Mathematics, University of Wisconsin, Madison, WI 53706.

ON THE ROUTH-HURWITZ PROBLEM

SH. STRELITZ

We consider the following well-known Routh-Hurwitz problem (R–H): to find necessary and sufficient conditions for all the zeros of a polynomial to lie in the left half of the complex plane (see for example [1]). Various solutions of this problem are known. In this paper we solve the R–H problem for a polynomial

$$(1) \qquad P(z) = z^n + a_1 z^{n-1} + \cdots + a_n$$

with real coefficients in new terms. It is obvious that the requirement of real coefficients in (1) is not essential, because if the real parts of all the zeros of a polynomial $P_0(z)$ with complex coefficients are negative, then the real part of every zero of the polynomial $P_0(z)\overline{P_0(\bar{z})}$, whose coefficients are already real, is also negative.

Besides the polynomial (1) we consider the polynomial

(2) $$Q(z)=z^n+b_1z^{m-1}+\cdots+b_m,\quad m=\frac{n(n-1)}{2},$$

the zeros of which are the sums and only the sums

$$z_i+z_j;\quad i<j,\quad j=1,2,\ldots,n,$$

where z_k, $k=1,2,\ldots,n$ are the roots of the equation $P(z)=0$. The coefficients b_k, $k=1,2,\ldots,m$ of the polynomial (2) are symmetric functions of z_k, $k=1,2,\ldots,n$, and can be found without the knowledge of the zeros of (1). In this paper we prove the following assertion.

THEOREM. *All the zeros of the polynomial $P(z)$ with real coefficients lie in the left half of the complex plane if and only if all the coefficients a_j of $P(z)$ and b_k of $Q(z)$ are positive, that is*:

$$a_j>0,\quad j=1,2,\ldots,n;$$
$$b_k>0,\quad k=1,2,\ldots,m,\quad m=n(n-1)/2.$$

Proof. The conditions are necessary. Let all the zeros $z_1,z_2,\ldots,z_n$ of (1) lie in the left half of the complex plane. Re $z_k<0$, $k=1,2,\ldots,n$. If $P(z_p)=0$, $z_p=\alpha+i\beta$, then also $P(\bar z_p)=0$, $\bar z_p=\alpha-i\beta$. Then in the expression

(3) $$(z-z_p)(z-\bar z_p)=z^2-2\alpha z+\alpha^2+\beta^2,$$

$-2\alpha>0$ and $\alpha^2+\beta^2>0$. Let z_q be a real zero and $z_q<0$. Then in the binomial $z-z_q$ we shall have $-z_q>0$. Suppose $\alpha_k+i\beta_k$, $\alpha_k-i\beta_k$, $k=1,2,\ldots,s$ are all the complex and γ_j, $j=1,2,\ldots,t$, $t=n-2s$ — all the real zeros of $P(z)$. Then

(4) $$P(z)=\prod_{k=1}^{n}(z-z_k)=\prod_{k=1}^{s}(z^2-2\alpha_kz+\alpha_k^2+\beta_k^2)\prod_{l=1}^{t}(z-\gamma_l).$$

After multiplying the factors in (4) we conclude according to the considerations above about the positiveness of $-\alpha_k$ and $-\gamma_k$ that all the coefficients in (1) are positive: $a_j>0$.

Furthermore, if Re $z_k<0$, $k=1,2,\ldots,n$, then all the zeros of $Q(z)$ lie in the left half of the complex plane and therefore, as we have proven, $b_k>0$, $k=1,2,\ldots,m$ too.

The conditions are sufficient. Let $\alpha+i\beta$ and $\alpha-i\beta$ be two zeros of $P(z)$ with $\alpha\neq0$ and $\beta\neq0$. Then 2α is a zero of $Q(z)$: $Q(2\alpha)=0$. But all the coefficients b_k of $Q(z)$ are positive. Thus all the real roots of the equation $Q(z)=0$ are negative and consequently all the complex roots of the equation $P(z)=0$ lie in the left half of the complex plane. Furthermore, all the real zeros of $P(z)$ lie in the left real axis, because $a_j>0$, $j=1,2,\ldots,n$.

The theorem is proved.

It seems that the conditions of this theorem can be successfully used for computer calculations. For this purpose we indicate below the relations between the coefficients b_k and a_j. Let z_k, $k=1,2,\ldots,n$ be the zeros of $P(z)$. Denote

$$\sigma_j=\sum_{k=1}^{n}z_k^j;\quad s_j=\frac{1}{2}\sum_{\substack{p,q=1\\p\neq q}}^{n}(z_p+z_q)^j,\quad j=0,1,2,3,\ldots.$$

The Newton relations give us:

(5)
$$\sigma_1+a_1=0,$$
$$\sigma_2+\sigma_1a_1+2a_2=0,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\sigma_n+\sigma_{n-1}a_1+2\sigma_{n-2}a_2+\cdots+na_n=0.$$

For $j > n$ we have further:

$$\sum_{k=1}^{n} z_k^{j-n}(z_k^n + a_1 z_k^{n-1} + \cdots + a_n) = \sigma_j + \sigma_{j-1}a_1 + \cdots + \sigma_{j-n} = 0.$$

Expanding both sides of the identity

$$\left(\sum_{k=1}^{n} e^{z_k t}\right)^2 = e^{2z_1 t} + e^{2z_2 t} + \cdots + e^{2z_k t} + \sum_{\substack{p,q=1\\ p\neq q}}^{n} e^{(z_p+z_q)t}$$

in Taylor series we get:

$$\left(\sum_{j=0}^{\infty} \frac{\sigma_j}{j!}\, t^j\right)^2 = \sum_{j=0}^{\infty} \frac{2^j\sigma_j}{j!}\, t^j + \sum_{j=0}^{\infty} \frac{2s_j}{j!} t^j.$$

Equating the coefficients of equal degrees of t we deduce the relations:

$$\sum_{p=0}^{j} \binom{p}{j} \sigma_p \sigma_{j-p} - 2^{j-1}\sigma_j = 2s_j.$$

After evaluating s_j we find b_j according to (5) (first replacing σ_j, a_j and n by a_j, b_j and m).

Reference

1. N. N. Meiman, and N. G. Chebotariov, The Routh-Hurwitz problem for polynomials and entire functions (Russian). Trudy Mat. Inst. ANSSSR 26, (1949) 1–331.

Department of Mathematics, University of Haifa, Mt. Carmel, Haifa 31 999, Israel.

SOME EQUATIONALLY COMPLETE ALGEBRAS

Hilda F. Duncan

The purpose of this note is to present elementary proofs of some of the theorems on equational completeness. These results were first published by Kalicki and Scott in 1955.

Many classes of algebras are described by a set of equations. If the set of elements of an algebra contains only one member, then every operation performed on this single element, or combinations of operations on this element, must of necessity produce this same element. This makes the single element algebra a trivial case in which every possible equation is true. We shall, therefore, consider in this discussion only sets of elements containing more than one member. A set of axioms will therefore be considered consistent if and only if there exists at least one algebra satisfying it that has at least two elements. An equation is derivable from a set of axioms if it is true in every algebra that satisfies the axioms.

Let us suppose then that we have two or more elements, an operation or operations and a set of equations. The axiom system thus described is equationally complete if any equation involving the elements and operations of the system is either derivable from the original set of equations or is not consistent with them. An algebra will be considered equationally complete if it satisfies a set of axioms that are equationally complete.

Let us now consider a very simple algebraic system with one unary operation. The result of applying this operation to an arbitrary element in our algebra will be designated by x'. Applying it twice will give $(x')'$ which we will abbreviate as x^2, and so on. Let us suppose that as our axiom system we have the single equation $x' = y'$, assumed true for all x and y. Any equation in this system which has only one unary operation must necessarily be of the form $x^p = y^q$ or $x^p = x^q$.

Let us consider $x^p = y^q$. If $p = q = 0$, we have $x = y$ which is a contradiction of the assumption that we have at least two distinct elements. Suppose $p = 0$ but $q \geqq 1$. Then we have $x = y^q = (y^{q-1})'$. Now denote y^{q-1} by z so that we have for all $x, x = z'$. But our axiom $x' = y'$ for all x and y implies that $x' = z'$. Therefore $x = x'$ for all x. Combining this result with our axiom $x' = y'$ we arrive at the conclusion $x = y$ which shows inconsistency. If $p \geqq 1$ and $q \geqq 1$ then $x^p = y^q$ becomes $(x^{p-1})' = (y^{q-1})'$ which follows immediately from our single axiom $x' = y'$ for all x and y.

Now suppose we have an equation of the form $x^p = x^q$. If $p = q = 0$ we have a trivial identity $x = x$. If $p = 0$ and $q \geqq 1$ we have $x = (x^{q-1})'$. If as above, $x^{q-1} = z$ we have for all x, $x = z'$ which leads to a contradiction as before. If $p \geqq 1$; $q \geqq 1$, $x^p = x^q$ may be written as $(x^{p-1})' = (x^{q-1})'$ which follows immediately from $x' = y'$.

Thus every possible equation in a system with one unary operation is derivable from, or inconsistent with, the equation $x' = y'$. Thus the system having this equation as its sole axiom is equationally complete.

It can be verified in a similar fashion that the systems with one unary operation and the single axiom $x = x'$ are also equationally complete.

To show that the systems characterized by the equations $x' = y'$ and $x = x'$ are the only equationally complete systems having a single unary operation we shall look at all possible equations and combinations of such equations that may occur in such a system. We shall show that every such equation or combination of equations is either inconsistent in itself or may be derived from at least one of the above equations.

As noted previously, any equation involving a single unary operation will of necessity have the form $x^p = y^q$ or $x^p = x^q$. In the first case, if either p or q is equal to zero, we have an equation of the form $x = y^q$ which is true for all x and y. Therefore, we also have $y = y^q$ and thus $x = y$. So we conclude that we may have $x^p = y^q$ only if $p \geqq 1$. But this case is directly implied by $x' = y'$.

Consider now $x^p = x^q$. If $p = q = 0$, we have the trivial case $x = x$. Suppose then that at least one value, say $q \geqq 1$. The resulting equation would then follow from $x = x'$. If we also have $p \geqq 1$, then it would also follow from $x' = y'$. We thus have shown three types of equations to be possible, namely $x^p = y^q (p \geqq 1,\ q \geqq 1)$, $x^p = x^q (p \geqq 1,\ q \geqq 1)$ and $x = x^q (q \geqq 1)$. The first two are implied by $x' = y'$ and the second two by $x = x'$. The only problem that might arise then would be from a combination of the first and the third. Let us suppose that we have $x^p = y^q (p \geqq 1,\ q \geqq 1)$ and $x = x^n (n \geqq 1)$. We may assume that p is less than n, say $n = p + r\ (r > 0)$, for if not, we may substitute into $x^p = y^q$ using $x = x^n$ until we arrive at a value for p that is less than n. Then we have $x = x^n = (x^p)^r = (y^q)^r = y^{q+r}$ which, as shown before, leads to the contradiction $x = y$.

Let us now look at algebras that have one binary operation which we will denote by "$\cdot$". We will first show that the system having the equation $w \cdot x = y \cdot z$ as its sole axiom is an equationally complete system. Our assumption ($w \cdot x = y \cdot z$) is valid for all w, x, y and z if and only if for all x and y, $x \cdot y = c$ where c is some constant. Let us break down the equations that might occur in such an algebra into three types. First we have the ones in which the binary operation does not appear on either side of the equal sign. These would be of the form $x = x$ or $x = y$, the first true and second false. Second, the binary operation appears on only one side (say the right) of our equation, so that our equation is of the form $x =$ an expression containing the binary operation at least once. Note that the right hand side will then reduce to c, the constant referred to previously, so that $x = c$ for all x. But this is a contradiction. Lastly consider the case in which the binary operation occurs at least once on both sides of the equation. Then both sides reduce to c and we arrive at the obviously true identity $c = c$.

The algebras having the axioms $x = x \cdot y$ and $y = x \cdot y$ respectively can similarly be shown to be equationally complete.

Another equationally complete algebra with one binary operation is the semi-lattice described by the equations $x \cdot (y \cdot z) = (x \cdot y) \cdot z$, $x \cdot y = y \cdot x$ and $x \cdot x = x$.

Let $t_1 = t_2$ be an equation in our system. We can use the commutative, associative and idempotent axioms to get each variable down to at most one occurence on each side. Let the simplified equivalent equation be denoted by $t_1' = t_2'$. Then we have two cases. (This proof was indicated to me by Professor Hugo Ribeiro.)

In the first case let us assume that we have exactly the same variables in t_1' as in t_2'. By the associative and commutative laws $t_1' = t_2'$ is equivalent to an equation in which the sides are identical, and the equation $t_1 = t_2$ is true in all semi-lattices.

In the second case let us assume that there is at least one variable v that occurs in t_1' but not in t_2'. Replace v by x and all other variables by y. We shall then get either $x = y$ (i.e., inconsistency) or $x \cdot y = y$. If the latter case occurs we can go back to $t_1' = t_2'$ and replace v now by y and all other variables by x. We shall then get $y \cdot x = x$. But by the commutative property $x \cdot y = y \cdot x$, so we have $x = y$ which again shows inconsistency.

The final set of axioms with one binary operation that we shall consider are those of the p-groups. If p is a prime ($p \geqq 2$) then we shall define a p-group to be an algebra that satisfies the equations $x \cdot (y \cdot z) = (x \cdot y) \cdot z$, $x \cdot y = y \cdot x$ and $x^p \cdot y = y$, where x^2 will be interpreted as $x \cdot x$, $x^3 = x \cdot x \cdot x$, etc. We intend to show that such a system is equationally complete. Note first that $x^p \cdot y = y$ implies that $x^{mp} \cdot y = y$ $(m \geqq 1)$. Fix p and suppose that we have an arbitrary equation involving one binary operation and an arbitrary number of variables. Let r_1 be the sum of all the x exponents on the left and s_1 be the sum of all the x exponents on the right. Similarly for all the other variables that occur.

Now if for all r_i and s_i that occur in the given equation we have $|r_i - s_i| = np$ for some $n \geqq 0$, then we have that the equation is implied by the corresponding p-group. This follows since $|r_i - s_i| = np$ for all i implies that for each i, either $r_i = np + s_i$ or $s_i = np + r_i$. Assume the latter, then

$$x^{s_i} \cdot w = x^{np+r_i} \cdot w = x^{np} \cdot x^{r_i} \cdot w = x^{r_i} \cdot w.$$

Since this is true of all the variables in the equation, we would have upon multiplying

$$x^{r_1} \cdot w \cdot y^{r_2} \cdot w \cdot z^{r_3} \cdot w \cdots = x^{s_1} \cdot w \cdot y^{s_2} \cdot w \cdot z^{s_3} \cdot w \cdots$$

Using commutativity and associativity we have $w^m \cdot x^{r_1} \cdot y^{r_2} \cdot z^{r_3} \cdots = w^m \cdot x^{s_1} \cdot y^{s_2} \cdot z^{s_3} \cdots$ Now there exists an $n \geqq 1$ such that $np \geqq m$ and $m + c = np$ $(c \geqq 0)$. Multiply both sides by w^c and since $w^{m+c} = w^{np}$ we have $x^{r_1} \cdot y^{r_2} \cdot z^{r_3} \cdots = x^{s_1} \cdot y^{s_2} \cdot z^{s_3} \cdots$ which by commutativity and associativity can be rearranged to produce the original equation.

Now suppose it is not true that $|r_i - s_i| = np$ $(n \geqq 0)$ for all i. This would mean that given any prime p there would exist some variable, say x, with sums of exponents r and s where $|r - s| = a \neq np$ $(n \geqq 0)$. Thus $r = s + a$ or $s = r + a$. Let us assume the latter. Let R equal the sum of the exponents of all the variables on the left, and S equal the sum of all the exponents on the right. Then letting every variable be equal to x we have $x^R = x^S$. Now in the original equation replace x by x^2. Then the left hand sum of the exponents would become $R + r$ and the right hand sum would become $S + s = S + r + a$. Again letting everything be equal to x we have combining the above results $x^{R+r} = x^{S+r+a} = x^{R+r+a}$. Now there exists an $m \geqq 1$ such that $mp \geqq R + r$. Then $mp = (R + r) + c$ $(0 \leqq c)$. Thus $y = x^p \cdot y = x^{mp} \cdot y = x^{R+r+c} \cdot y = x^{R+r+a+c} \cdot y = x^{R+r+c} \cdot x^a \cdot y = x^a \cdot y$. Now if $a = 1$ we have $x \cdot y = y$, but this implies $y \cdot x = x$ and $x \cdot y = y \cdot x$ so $x = y$. If $a \neq 1$ then since $p \geqq 2$ and they are relatively prime we can assume that the larger, say p can be written as $p = na + b$ $(0 < b < a$ and $n \geqq 1)$ and we have $x^p = x^{na} \cdot x^b = x^b$. If $b > 1$, we can in turn write $a = mb + d$ $(0 < d < b$ and $m \geqq 1)$ since a and b would be relatively prime. This can be continued until we have a remainder of 1 which would give $x \cdot y = y$ which leads to the contradiction described above. Similar reasoning would lead to the same result if $a > p$.

We have that a constant algebra $(x \cdot y = c)$, a left hand algebra $(x \cdot y = x)$, a right hand algebra $(y \cdot x = x)$, a semi-lattice and a p-group are all equationally complete systems of algebras each having just one binary operation. It will now be shown that any associative algebra with one binary operation is equationally complete only if it is one of the algebras listed above.

First consider any associative algebra that is both commutative and idempotent. It will by definition be a semi-lattice.

Second, assume that we have associativity and idempotence but not commutativity. If we have an equation where the first variable on the left is the same as the first variable on the right it will be implied by the left hand algebra. On the other hand, if we have the last variable on the left, the same as the last variable on the right, then it will be implied by the right hand algebra. Suppose we have neither. Let us designate the last variable on the right by w. We will then consider two cases. First assume that the first variable on the right is also w. Then letting every other variable to be equal to x and using the idempotence property we have the following four possibilities: (1) $x = w$, (2) $x = w \cdot x \cdot w$, (3) $x \cdot w \cdot x = w$, (4) $x \cdot w \cdot x = w \cdot x \cdot w$. The first is an obvious contradiction. In the second if we multiply both sides on the right by x and use idempotence we have $x = w \cdot x$. Repeating the procedure on the left in the original equation we get $x = x \cdot w$. Thus $x \cdot w = w \cdot x$ which contradicts our assumption that the algebra is not commutative. Equation three will similarly lead to the same contradiction if it is so multiplied by w. In equation (4) again multiplying first on the right and then the left of both sides by w leads to the equations $x \cdot w = w \cdot x \cdot w$ and $w \cdot x = w \cdot x \cdot w$. Thus $x \cdot w = w \cdot x$.

Let us assume that the first variable on the right is not w. Call it z. As before, let everything except w be equal to x. We have the following possibilities:

(1) $x = x \cdot w$, (2) $w \cdot x = x \cdot w$, (3) $x \cdot w \cdot x = x \cdot w$.

Now in the same equation let everything except z be equal to x. We then may have, (1′) $x = z \cdot x$, (2′) $x \cdot z = z \cdot x$, (3′) $x \cdot z \cdot x = z \cdot x$. Consider case (1). Then (1′), (2′) or (3′) will also hold: Case (2′) gives an immediate contradiction. Suppose (1′) holds. Then letting $z = w$ we have $x = w \cdot x$ which combined with (1) contradicts the noncommutativity. Any of the other combinations can be shown by methods similar to the preceding to also produce contradictions.

We have shown that without commutativity an individual equation must begin or end with the same variable on each side, and further, that all those beginning with the same variable are implied by the left hand algebra and those ending with the same variable by the right hand algebra. Suppose we have a pair of equations where there is one of each kind. Assume equation (1) begins on each side with an x but one side, say the right, ends in a w which is not the last variable on the left. Correspondingly in equation (2), assume that both sides end in an x but one side begins in a w that is not the first variable on the other side. In each case letting everything except w be equal to x and using idempotence an equation of type (1) would reduce to (a) $x = x \cdot w$ or (b) $x \cdot w \cdot x = x \cdot w$. Equation (2) would reduce to (c) $x = w \cdot x$ or (d) $x \cdot w \cdot x = w \cdot x$. Clearly the combinations (a) and (c) or (b) and (d) lead to contradictions. Suppose then we have (a) and (d). Using (a) $x = x \cdot w$ to substitute in (d) we have $x = w \cdot x$ which combined with (a) produces the contradiction. Similarly for (b) and (c).

Suppose we have an associative algebra that is not idempotent. If we have an equation or system of equations each of which involves the binary operation at least once on each side it will be implied by the constant algebra. Suppose we have a non-trival equation that contains a single variable on one of the sides. From our previous work on p-groups we know that if r_i equals the sum of all the exponents of a given variable on the left side of an equation, and s_i equals the sum of all the exponents of the corresponding variable on the right side, then the equation is implied by a p-group if there exists a prime p ($p \geqq 2$) such that $|r_i - s_i| = np$ (n an integer $\geqq 0$) for all i. If x is our single variable, say on the left of the equation, then if s_1 equals the sum of all the exponents of the x's on the right we know that $s_1 \geqq 1$, for otherwise we would have x equal to an expression not containing x which would lead to an immediate contradiction. Thus we would have $s_1 - 1 = np$ and $s_i = np$ for all other i. Suppose however there is no prime p for which this is true in some given equation containing only the single variable x on the left. Then there exist integers $r_1, r_2, \ldots, r_n$ such that $r_1(s_1 - 1) + r_2 s_2 + \cdots + r_n s_n = 1$. If $s = (s_1 - 1) + s_2 + \cdots + s_n$ then letting every variable be equal to x we have $x = x^s \cdot x$. Replacing the x's in the original equation by x^2 and then again letting all the variables be equal to x, we have $x^2 = x^{s+s_1}x = x^s \cdot x^{s_1-1} \cdot x^2 = x^{s_1-1} \cdot x^2$. Repeating this with each of the other variables, we get

$x = x^{s+s_i}x = x^s \cdot x^{s_i}x = x^{s_i} \cdot x$ or $x^2 = x^{s_i} \cdot x^2$ for all $i \neq 1$. Choose an m so that $m + r_i$ is positive for all i. Then

$$x^2 = x^{(m+r_1)(s_1-1)} \cdot x^{(m+r_2)s_2} \cdots x^{(m+r_n)s_n} x^2 = x^{m(s_1-1)} \cdot x^{ms_2} \cdots x^{ms_n} \cdot x^1 \cdot x^2 = x^{ms} \cdot x^3 = x^3.$$

Therefore $x^2 = x^{s+1} = x$ which is a contradiction.

Now suppose we have a system of equations where at least one of the non-trivial equations contains a single variable, say x on one side. If there exists a prime p such that in every equation we have $|r_i - s_i| = np$ $(n \geqq 0)$ for all i then all of the equations will be implied by the corresponding p group. Suppose this is not the case. Then as before from the single variable equation we have $x = x^s \cdot x$ where $s \geqq 1$. Let $x^{r_1} \cdot y^{r_2} \cdots z^{r_n} = x^{s_1} \cdot y^{s_2} \cdots z^{s_n}$ be any other equation in the system. Letting r' and s' be the sum of the left and right hand exponents respectively, we have upon letting all variables be x, $x^{r'} = x^{s'}$. Suppose for a given i we have $|r_i - s_i| = r_i - s_i$. Then upon replacing the corresponding variable by x^2 and all others by x we have

$$x^{s'+s_i} = x^{r'+r_i} = x^{s'+s_i+(r_i-s_i)}.$$

Now there exists a k such that $ks = m(s' + s_i) + c$ $(m \geqq 1)$. Therefore

$$x = x^s \cdot x = x^{ks} \cdot x = x^{m(s'+s_i)+c} \cdot x = x^{(m-1)(s'+s_i)+c} \cdot x^{s'+s_i+(r_i-s_i)} \cdot x = x^{m(s'+s_i)+c} \cdot x^{(r_i-s_i)} \cdot x = x^{r_i-s_i} \cdot x.$$

Similarly if $|r_i - s_i| = s_i - r_i$ we have $x = x^{s_i-r_i} \cdot x$. Thus in all equations and for all i we have $x = x^{|r_i-s_i|} \cdot x$. If we have u equations then there exist integers $r_i, t_i, \ldots, w_i$ such that

$$r_1|r_1^{(1)} - s_1^{(1)}| + \cdots + t_1|r_1^{(2)} - s_1^{(2)}| + \cdots + w_n|r_n^{(u)} - s_n^{(u)}| = 1,$$

where the superscript indicates the equation and the subscript the variable. Choose g so that $g + r_i$, $g + t_i$, $\cdots$, $g + w_i$ are positive for all i. Then

$$x = x^{(g+r_1)|r_1^1-s_1^1|} \cdots x^{(g+t_1)|r_1^2-s_1^2|} \cdots x^{(g+w_n)|r_n^u-s_n^u|} \cdot x$$
$$= x^{g|r_1^1-s_1^1|} \cdots x^{g|r_n^u-s_n^u|} \cdot x^1 \cdot x = x^2$$

which is of course a contradiction.

Reference

1. Jan Kalicki and Dana Scott, Equational completeness of abstract algebras, Indag. Math., 17 (1955) 650–659.

DEPARTMENT OF MATHEMATICS, PENNSYLVANIA STATE UNIVERSITY, DELAWARE COUNTY CAMPUS, 25 YEARSLEY MILL ROAD, MEDIA, PA 19063.

POLYNOMIALS AND INFINITE SUBFIELDS

ARTHUR CHARLESWORTH

Throughout this paper F denotes an infinite subfield of a field K.

Let $f(x)$ be a member of the polynomial ring $K[x]$ having the property that $f(F) \subseteq F$. It is not hard to see that $f(x)$ must then belong to $F[x]$. For let n be the degree of $f(x)$ and recall that there is exactly one polynomial of degree n or less over an infinite field which takes on given values at $n + 1$ elements of the field. (See [1], page 59.) Let $a_1, a_2, \ldots, a_{n+1}$ be distinct elements of F and let $L(x)$ be the unique polynomial of degree n or less in $F[x]$ such that

$$L(a_1) = f(a_1),\ L(a_2) = f(a_2), \ldots, L(a_{n+1}) = f(a_{n+1}).$$

Since $f(x)$ is a polynomial of degree n in $K[x]$ which agrees with $L(x)$ at these $n+1$ elements, and since $L(x)\in K[x]$, we have $L(x)=f(x)$ so $f(x)$ is in $F[x]$.

Thus we have shown

THEOREM 1. *Let F be an infinite subfield of K and let $f(x)$ belong to $K[x]$. If $f(F)\subseteq F$, then $f(x)$ belongs to $F[x]$.*

The purpose of this note is to extend Theorem 1 to an analogous statement about rational functions over K.

Let $f(x)$ and $g(x)$ belong to $K[x]$ and let $(f/g)(F)$ denote the set $\{f(a)/g(a) \mid a\in F, g(a)\neq 0\}$; we wish to show that if $(f/g)(F)\subseteq F$, then $f(x)$ and $g(x)$ must belong to $F[x]$. To have any hope of success, we clearly must impose several minor conditions on $f(x)$ and $g(x)$:

(i) Neither $f(x)$ nor $g(x)$ is the zero polynomial.

(ii) $f(x)$ and $g(x)$ are relatively prime in $K[x]$.

(iii) It is not possible to write the coefficients of both $f(x)$ and $g(x)$ as a fixed element of $K-F$ times elements of F.

The following theorem asserts that we need place no further conditions on $f(x)$ and $g(x)$.

THEOREM 2. *Let F be an infinite subfield of K and let $f(x)/g(x)$ belong to $K(x)$, where $f(x)$ and $g(x)$ satisfy the above conditions. If $(f/g)(F)\subseteq F$, then $f(x)$ and $g(x)$ belong to $F[x]$.*

Proof. Select a minimal list $e_1,\ldots,e_m$ of elements of K which are linearly independent over F and such that the coefficients of $f(x)$ and $g(x)$ can each be written in the form $c_1e_1+\cdots+c_me_m$ for some choice of c's in F. Then $f(x)=p_1(x)e_1+\cdots+p_m(x)e_m$ and $g(x)=q_1(x)e_1+\cdots+q_m(x)e_m$, where the $p(x)$'s and $q(x)$'s are in $F[x]$. If $m=1$, condition (iii) forces e_1 to be in F so $f(x)$ and $g(x)$ are in $F[x]$ and we are done; thus we may assume that $m\geqq 2$.

By condition (i) we may choose an infinite subset A of F such that $f(a)\neq 0$ and $g(a)\neq 0$ for all a in A. For each a in A let b_a in F be such that $f(a)/g(a)=b_a$; we can then rewrite $f(a)=b_ag(a)$ in the form

$$p_1(a)e_1+\cdots+p_m(a)e_m=b_a[q_1(a)e_1+\cdots+q_m(a)e_m],$$

where the $p(a)$'s and $q(a)$'s are elements of F. We thus have

$$p_1(a)=b_aq_1(a),\ p_2(a)=b_aq_2(a),\ldots,p_m(a)=b_aq_m(a)$$

so

$$\frac{p_1(a)}{q_1(a)}=\frac{p_2(a)}{q_2(a)}=\cdots=\frac{p_m(a)}{q_m(a)}. \tag{1}$$

Since each $q(x)$ is a polynomial and the list $e_1,\ldots,e_m$ was chosen to be minimal, we know that for all but finitely many a's in A each of $q_1(a),\ldots,q_m(a)$ is nonzero. Thus there is an infinite subset A^* of A on which all equations in (1) make sense. Notice that by cross multiplication

$$\frac{p_1(a)}{q_1(a)}=\frac{p_i(a)}{q_i(a)}\ \text{ for all } a \text{ in } A^* \text{ implies }\ \frac{p_1(x)}{q_1(x)}=\frac{p_i(x)}{q_i(x)}.$$

Let us denote $p_1(x)/q_1(x)$ by $R(x)$. Then

$$f(x)=q_1(x)R(x)e_1+q_2(x)R(x)e_2+\cdots+q_m(x)R(x)e_m$$

and, since

$$g(x)=q_1(x)e_1+q_2(x)e_2+\cdots+q_m(x)e_m,$$

we have $f(x)/g(x)=R(x)$ and thus $f(x)/g(x)=f_1(x)/g_1(x)$, where $f_1(x)/g_1(x)$ is $R(x)$ written in lowest terms over F. Thus $f(x)\cdot g_1(x)=g(x)\cdot f_1(x)$. Now $f(x)$ and $g(x)$ are relatively prime in $K[x]$

(by condition (ii)) and $f_1(x)$ and $g_1(x)$ are relatively prime in $K[x]$ (by [1], page 73, problem 17) so by unique factorization $f(x) = c_1 f_1(x)$ and $g(x) = c_2 g_1(x)$, for some constants c_1 and c_2 in K.

We conclude the proof by observing that c_1 and c_2 must be in F. Let a be any element of A^*. Since $f(a)/g(a) = (c_1/c_2)\cdot f_1(a)/g_1(a)$ and since both $f(a)/g(a)$ and $f_1(a)/g_1(a)$ are in F, we know that c_1/c_2 is in F. Now $f(x) = c_2(c_1/c_2)f_1(x)$ and $g(x) = c_2 g_1(x)$ so condition (iii) assures us that c_2 (and hence c_1) must belong to F. $\square$

We close with an example to show why it is important that F be infinite throughout this note. Let $\mathbf{Z}_2$ be the integers mod 2, let $\{0, 1, a, b\}$ be the field with four elements, and let $f(x) = ax^2 + bx + 1$. Then $f(\mathbf{Z}_2) \subseteq \mathbf{Z}_2$ (since $a + b = 1$), but $f(x) \notin \mathbf{Z}_2[x]$.

Reference

1. Garrett Birkhoff and Saunders Mac Lane, A Survey of Modern Algebra, Third Edition, Macmillan, New York, 1965.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF RICHMOND, RICHMOND, VA 23173.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics and Statistics, The University of Calgary, Calgary, Alberta, Canada, T2N 1N4.

SEVERAL CONJECTURES ON COMMUTATIVITY IN ALGEBRAIC STRUCTURES

KEITH S. JOSEPH

Given the algebraic structure A, we define $\Pr(A)$ to be the probability that a pair of elements of A will commute with each other. The function Pr has been studied for both finite groups [2, 5, 6] rings [7] and compact groups [5].

We first summarize some of the known results. Although there are analogs for other algebraic structures we will state the results for the case of a finite group G:

(1) $\Pr(G) = k(G)/|G|$ where $k(G)$ is the number of conjugacy classes of G.

(2) If H is a subgroup of G, then $\Pr(H)/[G:H]^2 \leqq \Pr(G) \leqq \Pr(H)$.

(3) If H is a normal subgroup of G, then $\Pr(G) \leqq \Pr(H)\Pr(G/H)$.

(4) Suppose G is a non-commutative group, and p is the smallest prime divisor of $|G|$. If $\Pr(G) > 1/p$, then

(i) $|G'| = p$ where G' denotes the commutator subgroup of G.

(ii) G' is contained in the center $Z(G)$ of G.

(iii) $[G: Z(G)] = p^{2s}$ for some $s \geqq 1$.

(iv) $\Pr(G) = (1/p) + (p-1)/p^{2s+1}$.

(5) If G is non-commutative, then $\Pr(G) \leqq 5/8$.

Now let V denote the set of values of the function Pr. Suppose there is a sequence $v_n \in V$ such that $\lim_{n\to\infty} v_n = k$. The above results, and other results of a similar nature, suggest the following conjectures:

CONJECTURE 1: *k is a rational number.*
CONJECTURE 2: *$v_n > k$ for (almost) all n.*
CONJECTURE 3: *There is a group G such that $\Pr(G) = k$.*

Equivalently Conjecture 1 states that if k is irrational then there exists $\delta > 0$, such that the interval $(k - \delta, k + \delta)$ contains no elements of V. Conjecture 2 states that if k is rational then there exists a $\delta > 0$ such that $(k - \delta, k)$ contains no elements of V.

The largest value of k for which such a sequence exists is $k = \frac{1}{2}$. Take $v_n = \frac{1}{2} + 1/2^{2n+1} = \Pr(G_n)$ where G_n is an extra-special 2-group [4, p. 183] with $[G: Z(G)] = 2^{2n}$. There are no elements of V in the interval (7/16, 1/2). Furthermore, $\Pr(S_3) = \frac{1}{2}$ where S_3 is the symmetric group on three letters. So all three conjectures are confirmed for $k = \frac{1}{2}$.

Finally, if Conjecture 2 is true then the set V is seen to be a well-ordered set. A. W. Hales (personal communication) has posed the following question: What is the order type of the set V?

References

1. J. A. Ernest, Central intertwining numbers for representations of finite groups, Trans. Amer. Math. Soc., 99 (1961) 499–508.

2. P. Erdös and P. Turán, On some problems of a statistical group-theory, IV, Acta Math. Acad. Sci. Hung., 19 (1968) 413–435.

3. P. X. Gallagher, The number of conjugacy classes in a finite group, Math. Z., 118 (1970) 175–179.

4. D. Gorenstein, Finite Groups, Harper and Row, New York, 1968.

5. W. H. Gustafson, What is the probability that two group elements commute? this MONTHLY, 80 (1973) 1031–1034.

6. K. S. Joseph, Commutativity in non-Abelian groups, Ph.D. Thesis, University of California, Los Angeles, 1969.

7. D. MacHale, Commutativity in finite rings, this MONTHLY, 83 (1976) 30–32.

SCHOOL OF SCIENCE AND MATHEMATICS, METROPOLITAN STATE COLLEGE, DENVER, CO 80204.

CLASSROOM NOTES

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Department of Mathematics, University of Wisconsin, Madison, WI 53706.

A CONVERSE OF PYTHAGORAS' THEOREM

S. GUDDER AND D. STRAWTHER

Pythagoras' theorem is probably the oldest mathematical theorem of any importance. The Babylonians used Pythagoras' theorem as early as 1700 B.C.E., although it was not actually proved until about 500 B.C.E. by the ancient Greeks. This theorem and its generalizations have had a profound influence on almost all fields of mathematics. Although Pythagoras' theorem has been with us for almost 4000 years, nobody, as far as we know, has considered the converse that we present here.

(A related, less general result, using different methods is given in [**2**].) Only elementary results and methods of real analysis are used and the proof should be easily accessible to mathematics majors in their senior year of college. In fact, the only results needed are:

1. L'Hospital's rule.

2. A nondecreasing real function is continuous, except for a countable number of points [**1**; p. 158].

3. The only continuous solution of $f(x+y)=f(x)+f(y)$ is $f(x)=cx$ [**1**; p. 158].

Let R^2 be the Euclidean plane in which the distance of $x=(x_1,x_2)\in R^2$ to 0 (or *norm* of x) is defined as $\|x\|=(x_1^2+x_2^2)^{1/2}$. If we define the inner product $\langle x,y\rangle = x_1y_1+x_2y_2$ for $x=(x_1,x_2)$, $y=(y_1,y_2)$ then $\|x\|^2=\langle x,x\rangle$. We say that x is *orthogonal* to y (denoted $x\perp y$) iff $\langle x,y\rangle=0$. Pythagoras' theorem may be stated as follows: $\|x+y\|^2=\|x\|^2+\|y\|^2$ whenever $x\perp y$. One converse of this theorem would be: if $\|x+y\|^2=\|x\|^2+\|y\|^2$ then $x\perp y$. But this converse is trivial and not particularly important. A more profound converse would state that $x\mapsto\|x\|^2$ is essentially the only function that satisfies Pythagoras' theorem. That is, if $f: R^2\to R$ is a distance function, in some sense, which satisfies the Pythagorean condition (i) $f(x+y)=f(x)+f(y)$ whenever $x\perp y$, then $f(x)=c\|x\|^2$ for some $c\geqq 0$. It is clear that condition (i) alone is not sufficient to characterize f since any linear function satisfies (i). Thus an additional condition is needed. This is the condition that f be a distance function in some sense. What sense? A natural condition that all distance functions satisfy is positivity (ii) $f(x)\geqq 0$ for all $x\in R^2$.

THEOREM. *If $f: R^2\to R$ satisfies* (i) $f(x+y)=f(x)+f(y)$ *whenever $x\perp y$ and* (ii) $f(x)\geqq 0$ *for all $x\in R^2$, then $f(x)=c\|x\|^2$ for some $c\geqq 0$.*

Proof. Step 1: We first show that $f(\alpha y)\geqq f(y)$ for all $y\in R^2$, $\alpha\geqq 1$. Let $x\in R^2$ satisfy $\|x\|=1$, $x\perp y$. Let $\beta=(\alpha-1)^{1/2}\|y\|$ and define $z=(\alpha-1)y-\beta x$. Then $z\perp y+\beta x$ so that

$$f(\alpha y)=f(y+\beta x+z)=f(y+\beta x)+f(z)=f(y)+f(\beta x)+f(z)\geqq f(y).$$

Step 2: We now show that if $\|y\|>\|x\|$ then $f(y)\geqq f(x)$. From (i) $f(0)=0$ so the result holds for $x=0$ using (ii). We thus assume that $x\neq 0$. Let $\theta=\text{angle}(y,x)$. Applying L'Hospital's rule to $\lim_{h\to 0}(\log\cos h\theta)/h$ we conclude that $\lim_{n\to\infty}(\cos\theta/n)^n=1$. Hence there exists an n such that $\alpha\equiv\|y\|(\cos\theta/n)^n/\|x\|\geqq 1$. Define elements $y_0=\alpha x, y_1, y_2,\dots, y_n=y$ as follows: $\text{angle}(y_i,y_{i-1})=\theta/n$, $\|y_{i-1}\|=\|y_i\|\cos\theta/n$, $1\leqq i\leqq n$. (Draw a picture!) Notice that $(y_i-y_{i-1})\perp y_{i-1}$, $1\leqq i\leqq n$. By (i) and (ii)

$$f(y)=f[y_{n-1}+(y_n-y_{n-1})]=f(y_{n-1})+f(y_n-y_{n-1})\geqq f(y_{n-1}).$$

Continuing this process and finally applying Step 1 we have

$$f(y)\geqq f(y_{n-1})\geqq f(y_{n-2})\geqq\cdots\geqq f(\alpha x)\geqq f(x).$$

Step 3: Let $x\neq 0$. We show that if $f(\alpha_i x)\to f(x)$ as $\alpha_i\to 1$, then f is continuous at x. Let x_i be a sequence converging to x. Clearly we can assume $x_i\not\perp x$. Now let $s_i=\langle x_i,x\rangle\|x\|^{-2}x$ and let $r_i=\|x_i\|^2\langle x, x_i\rangle^{-1}x$. Then $(r_i-x_i)\perp x_i$, $(x_i-s_i)\perp s_i$, $i=1,2,\dots$. Hence

$$f(r_i)=f(r_i-x_i+x_i)\geqq f(x_i)=f(x_i-s_i+s_i)\geqq f(s_i).$$

Since the coefficients of x in s_i and r_i converge to 1 we have

$$\lim_{i\to\infty} f(r_i)=\lim_{i\to\infty} f(s_i)=f(x)\quad\text{so}\quad \lim_{i\to\infty} f(x_i)=f(x).$$

Step 4: Let $x\neq 0$ and $S=\{\lambda x:\lambda>0\}$. We show that f is continuous at every point of S except for a countable set $\hat{S}\subseteq S$. By Step 2, f restricted to S is nondecreasing (regarded as a function of λ). Hence f restricted to S is continuous except for a countable set $\hat{S}\subseteq S$. We now show that f itself (not

just f restricted to S) is continuous on $S-\hat{S}$. Let $y\in S-\hat{S}$. Then $f(\alpha_i y)\to f(y)$ as $\alpha_i\to 1$ so it follows from Step 3 that f is continuous at y.

Step 5: We show that if f is continuous at x and $\|y\|=\|x\|$, then $f(y)=f(x)$. For $x=0$, the result holds trivially, so assume $x\neq 0$. If $\lambda>1$, then $\|\lambda x\|>\|y\|$ so by Step 2, $f(\lambda x)\geqq f(y)$. Letting $\lambda\to 1$ we conclude that $f(x)\geqq f(y)$. Similarly, $f(x)\leqq f(y)$. So $f(x)=f(y)$.

Step 6: We show that if f is continuous at x and $\|y\|=\|x\|$, then f is continuous at y. Again we can assume $x\neq 0$. Let $y_i\to y$. As $\|y\|=\|x\|>0$, it is possible to find a sequence $a_i\in R$ such that $a_i\to 0$, $a_i>0$ and $\|y_i\|-a_i>0$ for i sufficiently large. Let $x_i=(\|y_i\|+a_i)x/\|y\|$ and $z_i=(\|y_i\|-a_i)x/\|y\|$. Then $\|x_i\|>\|y_i\|>\|z_i\|$ so by Step 2, $f(x_i)\geqq f(y_i)\geqq f(z_i)$. Now $x_i\to x$, $z_i\to x$ so by the continuity of f at x and Step 5 we have $f(y_i)\to f(x)=f(y)$. Hence f is continuous at y.

Step 7: It follows from Steps 4 and 6 that f is continuous in R^2 except for a countable number of circles centered at 0.

Step 8: We now show that f is continuous everywhere on R^2. Let $x\neq 0$ and pick $\alpha_i\to 1$. By Step 7 there exists $y\in R^2$ such that $x\perp y$ and f is continuous at $x+y$. Since $\alpha_i x+y\to x+y$ we have $f(\alpha_i x)+f(y)=f(\alpha_i x+y)\to f(x+y)=f(x)+f(y)$. Hence $f(\alpha_i x)\to f(x)$. By Step 3, f is continuous at x. To show f is continuous at 0, let $\|x_0\|=1$. Then there exists $z\in R^2$ such that $z\perp x_0$ and f is continuous at z. Since $(1/n)x_0+z\to z$ as $n\to\infty$ we have $f[(1/n)x_0]+f(z)=f[(1/n)x_0+z]\to f(z)$. Hence $f[(1/n)x_0]\to 0$ as $n\to\infty$. Now let $x_i\to 0$. Given $\varepsilon>0$, there exists an N such that $f[(1/N)x_0]<\varepsilon$. If i is sufficiently large, then $\|x_i\|<1/N$ and hence by Step 2, $0\leqq f(x_i)\leqq f[(1/N)x_0]<\varepsilon$. Thus f is continuous at 0.

Step 9: If $\|x\|=\|y\|$ then by Steps 5 and 8, $f(x)=f(y)$. It follows that the function $g:[0,\infty)\to[0,\infty)$ given by $g(\lambda)=f(x)$ where $\|x\|=\lambda$ is well defined. Also g is continuous since f is. If $\lambda,\mu\in[0,\infty)$, let x and y be orthogonal and satisfy $\|x\|=\lambda$, $\|y\|=\mu$. Then

$$g(\lambda)+g(\mu)=f(x)+f(y)=f(x+y)=g(\|x+y\|)=g[(\lambda^2+\mu^2)^{1/2}].$$

Making a change of variables $\alpha=\lambda^2$, $\beta=\mu^2$ and letting $h(w)=g(w^{1/2})$ we see that $h:[0,\infty)\to[0,\infty)$ is continuous and satisfies $h(\alpha)+h(\beta)=h(\alpha+\beta)$ for all $\alpha,\beta\in[0,\infty)$. Now the only function with these properties is $h(\alpha)=c\alpha$ for some $c\geqq 0$. Hence $f(x)=g(\|x\|)=h(\|x\|^2)=c\|x\|^2$. Q.E.D.

The above proof can be easily altered so that the theorem holds for any inner product space of dimension at least two.

References

1. R. Bartle, The Elements of Real Analysis, Wiley, New York, 1964.
2. K. Sundaresan, Orthogonality and nonlinear functionals on Banach space, Proc. A.M.S., 34 (1972) 187–190.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF DENVER, DENVER, CO 80210.

A COMMENT ON UNIONS OF SIGMA-FIELDS

ALLEN BROUGHTON AND BARTHEL W. HUFF

It is well known that if $\mathcal{A}_1,\mathcal{A}_2,\ldots$ is a denumerable sequence of sigma-fields of subsets of S with $\mathcal{A}_1\subset\mathcal{A}_2\subset\cdots$ (where the inclusion is proper), then $\bigcup_{n=1}^{\infty}\mathcal{A}_n$ need not be a sigma-field. This fact appears as a standard exercise in such texts as **[1]**, p. 4 and **[2]**, p. 19. In this note, we show that a much stronger statement is possible; such a union can *never* be a sigma-field.

LEMMA. *Let $\mathcal{A}_1,\mathcal{A}_2,\ldots$ be a denumerable sequence of sigma-fields with $\mathcal{A}_1\subset\mathcal{A}_2\subset\cdots$. Then (we may assume that) there exists a sequence of disjoint sets $F_1,F_2,\ldots$ such that $F_k\in\mathcal{A}_{k+1}\setminus\mathcal{A}_k$.*

Proof: Without loss of generality, we may assume that $\mathcal{A}_1\neq\{\varnothing,S\}$; i.e., there exists $B\in\mathcal{A}_1$ such

that $\varnothing \subset B \subset S$. Now suppose that for some n

$$B \cap \mathscr{A}_{n+1} = B \cap \mathscr{A}_n$$

and

$$\tilde{B} \cap \mathscr{A}_{n+1} = \tilde{B} \cap \mathscr{A}_n = \{\tilde{B} \cap C \mid C \in \mathscr{A}_n\}.$$

Then if $X \in \mathscr{A}_{n+1} \setminus \mathscr{A}_n$, we see that $B \cap X \in B \cap \mathscr{A}_{n+1} = B \cap \mathscr{A}_n \subseteq \mathscr{A}_n$ and $\tilde{B} \cap X \in \tilde{B} \cap \mathscr{A}_{n+1} = \tilde{B} \cap \mathscr{A}_n \subseteq \mathscr{A}_n$, implying that $(B \cap X) \cup (\tilde{B} \cap X) = X \in \mathscr{A}_n$ and we have a contradiction. Thus there must exist a set $E \in \mathscr{A}_1$ (where $E = B$ or $E = \tilde{B}$) such that $E \cap \mathscr{A}_{n+1} \setminus E \cap \mathscr{A}_n \neq \varnothing$ infinitely often.

Let $n_1, n_2, \dots$ be a sequence of indices such that $E \cap \mathscr{A}_{n_{k+1}} \setminus E \cap \mathscr{A}_{n_k} \neq \varnothing$. Now set $\mathscr{C}_k = E \cap \mathscr{A}_{n_k}$. Then $\mathscr{C}_1 \subset \mathscr{C}_2 \subset \cdots$ is a denumerable sequence of sigma-fields of subsets of E. Thus, by the preceding argument, there exists a set $E_1 \in \mathscr{C}_1$ such that $\varnothing \subset E_1 \subset E$ and $E_1 \cap \mathscr{C}_{k+1} \setminus E_1 \cap \mathscr{C}_k \neq \varnothing$ infinitely often.

Continuing inductively, we obtain a subsequence $\mathscr{A}_{j_1}, \mathscr{A}_{j_2}, \dots$ of sigma-fields and a sequence $E_1 \supset E_2 \supset \cdots$ of sets such that $E_k \in \mathscr{A}_{j_k}$ and

$$E_{k+1} \in E_k \cap \mathscr{A}_{j_{k+1}} \setminus E_k \cap \mathscr{A}_{j_k}.$$

Setting $F_k = E_k \setminus E_{k+1}$ and restricting our attention to the subsequence of sigma-fields, we verify the lemma. ∎

THEOREM. *Let $\mathscr{A}_1, \mathscr{A}_2, \dots$ be a denumerable sequence of sigma-fields with $\mathscr{A}_1 \subset \mathscr{A}_2 \subset \cdots$. Then $\bigcup_{n=1}^{\infty} \mathscr{A}_n$ cannot be a sigma-field.*

Proof: We identify the set F_k of the lemma with $\{k\}$ and assume that we are dealing with sigma-fields on the set of positive integers such that $\{k\} \in \mathscr{A}_{k+1} \setminus \mathscr{A}_k$.

Let B_n be the smallest set in $\mathscr{A}_n$ containing n. Then $n \in B_n \subseteq \{n, n+1, n+2, \dots\}$, $B_n \neq \{n\}$, and if $m \in B_n$ then $B_m \subseteq B_n$ (since $m \in B_n \cap B_m \in \mathscr{A}_m$).

Define $n_1 = 1$ and having chosen n_k select $n_{k+1} \in B_{n_k}$ such that $n_{k+1} \neq n_k$. Then $B_{n_1} \supset B_{n_2} \supset \cdots$. Let $E = \{n_2, n_4, n_6, \dots\}$. If $\cup \mathscr{A}_n$ is a sigma-field, then $E \in \mathscr{A}_n$ for some n and thus $E \in \mathscr{A}_{n_{2k}}$ for some k. But then $\{n_{2k}, n_{2k+2}, \dots\} \in \mathscr{A}_{n_{2k}}$ and this implies that $B_{n_{2k}} \subseteq \{n_{2k}, n_{2k+2}, \dots\}$. This is a contradiction since $n_{2k+1} \in B_{n_{2k}}$. ∎

The authors thank the referee for a shorter proof of the Theorem.

References

1. H G. Tucker, A Graduate Course in Probability, Academic Press, New York, 1967.
2. K. L. Chung, A Course in Probability Theory, Harcourt, Brace and World, New York, 1968.

DEPARTMENT OF MATHEMATICS, QUEEN'S UNIVERSITY, KINGSTON, ONT. K7L 3N6, CANADA.

A NOTE ON MONOGENIC BAIRE MEASURES

ALVIN F. MARTIN

This note presents a simple counterexample which answers in the negative two questions posed by S. K. Berberian in [**1**, p. 233]. Specifically, we exhibit a locally compact Hausdorff space X satisfying each of the following: (1) X has a monogenic Baire measure which is not completion regular; and (2) every Baire measure on X is monogenic, but not every Borel set is a Baire set. This example is also presented in [**2**, p. 106], in connection with (2). However, the method of proof used there is more complicated, and (1) is not considered.

The terminology is that of [**1**] and [**3**]. The Baire sets of a locally compact Hausdorff space X are

the elements of the σ-ring generated by the compact G_δ subsets of X; the Borel sets are the elements of the σ-ring generated by the compact subsets of X. A Baire measure on X is monogenic if it has a unique Borel extension. It is not hard to show (see [**1**, p. 214] or [**3**, p. 229]) that every Baire measure ν has a unique Borel extension μ which is regular, meaning that for each Borel set E

$$\mu(E) = \mathrm{GLB}\{\mu(U) \colon E \subset U, U \in \mathcal{U}\} = \mathrm{LUB}\{\mu(C) \colon C \subset E, C \in \mathcal{C}\},$$

where $\mathcal{U}$ is the class of open Borel sets of X, and $\mathcal{C}$ is the class of compact subsets of X. A Baire measure ν is completion regular if for every Borel set E there are Baire sets G and F such that $G \subset E \subset F$ and $\nu(F-G)=0$. Clearly, completion regularity implies monogenicity.

Let X_0 be a discrete space of uncountable but weakly accessible cardinality. (Or simply, let X_0 have cardinality $\aleph_1$.) Let $X = X_0 \cup \{e\}$ be the one point compactification of X_0. Then every subset of X is a Borel set, but only the countable subsets of X_0 and their complements are Baire sets. In particular, $\{e\}$ is not a Baire set. Define a Baire measure ν on X by letting $\nu(F)$ be 1 or 0 according as e does or does not belong to F. Since there do not exist Baire sets F and G with $G \subset \{e\} \subset F$ and $\nu(F-G)=0$, ν is not completion regular.

However, ν is monogenic. Its regular Borel extension μ is defined by taking $\mu(E)$ to be 1 or 0 according as e does or does not belong to E. Suppose μ' were a different Borel extension. Then $\mu' | X_0$ would be a measure on the σ-ring of all subsets of X_0 such that $\mu'(X_0) > 0$, but $\mu'(\{x\}) = 0$ for each $x \in X_0$. This would contradict a famous theorem of Ulam [**4**].

Ulam's theorem implies that every Borel measure on X is concentrated on a countable subset. From this it follows easily that every Borel measure on X is regular. Thus every Borel measure on X is monogenic, since regular Borel extensions are unique.

I wish to thank the referee for pointing out the reference [**2**].

References

1. S. K. Berberian, Measure and Integration, Macmillan, New York, 1965.

2. R. J. Gardner, The regularity of Borel measures and Borel measure-compactness, Proc. London Math. Soc., (3) 30 (1975) 95–113.

3. P. R. Halmos, Measure Theory, Van Nostrand, Princeton, N.J., 1950.

4. S. Ulam, Zur Masstheorie in der allgemeine Mengenlehre, Fund. Math., 16 (1930) 140–150.

DEPARTMENT OF MATHEMATICS, YALE UNIVERSITY, NEW HAVEN, CT 06520.

MATHEMATICAL EDUCATION

EDITED BY PAUL T. MIELKE AND SHIRLEY HILL

Material for this Department should be sent to Paul T. Mielke, Department of Mathematics, Wabash College, Crawfordsville, IN 47933, or to Shirley Hill, Department of Mathematics, University of Missouri, Kansas City, MO 64110.

MODULAR APPLIED MATHEMATICS FOR BEGINNING STUDENTS

ERVIN Y. RODIN

Introduction. There seems to be a growing interest in the application of mathematics to nonconventional fields, and, concurrently, a general interest in Applied Mathematics. Thus the question of how and when to teach Applied Mathematics, and even the question of what to teach, arises more and more often.

TRENDS IN CONTENT PROGRAMS FOR PRESERVICE SECONDARY MATHEMATICS TEACHERS

CARL S. JOHNSON AND JACKSON A. BYARS

This is a report of a status study that was conducted in 1974 on the preparation of secondary school mathematics teachers in the United States [**4**]. Data for the study were gathered from questionnaires which were sent to the chairperson of the mathematics department in 749 institutions of higher learning in the United States. All of these institutions of higher learning were listed in *The American Association of Colleges for Teacher Education Directory* [**1**]. They were also listed in the ninth edition of *The College Blue Book* as a four-year college or university by 1959 [**2**].

Four hundred forty-eight (60 percent) questionnaires were returned and data from 418 of the questionnaires were used in the analysis. Thirty of the questionnaires were not used for reasons such as the following: (1) the institution prepared only elementary school teachers; (2) the institution offered mathematics courses only to graduate students; and (3) the questionnaire was inappropriate to their particular institution.

The writers examined the characteristics of the respondent institutions and determined that there were no systematic differences between them and non-respondent institutions in terms of the following variables: geographic location, the highest degree offered by the mathematics department, and number of semester hours required for a mathematics major.

The 749 institutions of higher learning were divided into three groups. Group I included 489 institutions which offered the bachelor's degree as the highest degree in mathematics. Group II consisted of 157 institutions which offered the master's degree as the highest degree in mathematics. Group III included 103 institutions which offered the doctorate as the highest degree in mathematics. Included in the 418 usable questionnaires were 249 from Group I, 105 from Group II, and 64 from Group III.

Findings. One section of the questionnaire asked for the number of semester hours required for a major for teachers of junior high and senior high school mathematics. The mean number of semester hours of mathematics required by institutions of higher learning which offered a specific preservice junior high program was 31.42 semester hours, while 33.28 semester hours were required for a preservice senior high school mathematics major. Table 1 gives a summary of these requirements.

Information on content courses in mathematics required of or frequently taken by preservice secondary school mathematics teachers comprised another section of the questionnaire. The mean number of semester hours required of those courses listed on the questionnaire were: elementary calculus (10.83), modern algebra (2.72), geometry (2.45), linear algebra (2.05), probability and statistics (1.56), computer science (0.67), history of mathematics (0.23), foundations of mathematics (0.85), set theory and logic (0.38), introduction to analysis (0.72), number theory (0.23), topology (0.10), differential equations (0.47), the real number system (0.15), advanced calculus (0.79), numerical analysis (0.79), applied mathematics (0.05), real variables (0.14), complex variables (0.04), and other required courses (0.60).

The CUPM Level III recommendations for secondary school mathematics teachers are: three courses in calculus, one course in real analysis, two courses in algebra, two courses in probability and statistics, two courses in geometry, one course in applications, and experience with applications of computing [**3**]. While these recommendations were designed primarily for prospective secondary school teachers, they were also constructed with the idea of maintaining a comparability of standards between prospective teachers and prospective graduate students who had a major in mathematical science. The recommendations state that the program for teachers of secondary school mathematics should be identical to the one which is offered to other mathematics majors, with the exception of a few courses designed specifically for the secondary school mathematics teacher.

TABLE 1

Number of semester hours of mathematics required for teachers of Junior and Senior High School mathematics in institutions of higher learning in the United States in 1974.

Semester Hours	Institutions Which Offer a Specific Jr. High Program	Institutions Which Offer a Specific Sr. High Program	Institutions Which Offer a Secondary Program Not Specific to Levels
	Number	Number	Number
12–15	4	0	0
16–19	14	0	0
20–23	14	3	1
24–27	40	29	6
28–31	83	95	42
32–35	77	93	28
36–39	39	44	20
40–43	27	31	12
44–47	8	9	2
48–51	1	3	0
Totals	307	307	111
Means	31.42	33.28	33.25
Medians	31.43	32.64	32.43

The recommendations also encouraged use of the computer in mathematics courses whenever possible. This recommendation was made in light of the fact that more secondary schools are beginning to utilize the computer. It is further recommended that added emphasis be placed on applications in mathematics courses. This could be done by utilizing the concept of building a mathematical model in order to lay a concrete foundation on which to build a conceptual framework for theoretical and abstract thinking.

It is readily apparent that high school geometry continues to change. Therefore, the CUPM recommends that teachers should be prepared to teach geometry from either the modern Euclidean approach or from the new algebraic point of view.

Assuming that a college mathematics course is worth three semester hours of credit, the CUPM recommendations are being met fairly well in calculus, analysis (defined here as those courses in introduction to analysis, differential equations, advanced calculus, real variables, complex variables, and numerical analysis), and algebra (defined here as those courses in modern algebra and linear algebra). However, many institutions of higher learning are failing to meet the CUPM recommendations in their requirements by at least four semester hours in probability and statistics, three semester hours in geometry, and two semester hours in both computer science and applications (defined here as applied mathematics). These comparisons are shown in Table 2.

Although the above discussion indicates a program might be falling short in meeting several requirements, it must be kept in mind that the typical program, based on the averages, shows approximately 12 semester hours of electives. It would seem reasonable to believe that the typical student, completing a program of study in the respondent institutions, would select at least some of his electives in such a manner as to have a program corresponding more closely to the CUPM recommendations than the averages indicate.

TABLE 2

Mathematics courses required and trends observed in institutions of higher learning in the United States

Subject Area	Mean no. of Semester Hours Required	No. of Institutions Indicating an Increase Since 1960 as an Important Change	No. of Institutions Indicating a Need to Increase as Important	CUPM Recommendations	
				No. of Semester Hours	Met
Calculus	10.83	2	0	9	Yes
Analysis*	2.95	80	5	3	Yes
Modern Algebra	2.72	105	10	3	Yes
Linear Algebra	2.05	111	1	3	Yes
Probability & Statistics	1.56	95	28	6	No
Geometry	2.45	117	38	6	No
Computer Science	0.67	108	50	3	No
Applications	0.05	6	47	3	No

* Analysis is defined here as courses in introduction to analysis, differential equations, advanced calculus, real variables, complex variables, and numerical analysis.

When a closer look is taken at what has happened in course offerings in mathematics in the past fifteen years, one can see that much progress has been made by institutions of higher learning in meeting the CUPM recommendations. Information from Table 2 shows the number of schools which have increased their mathematics course offerings since 1960 as follows: analysis (80), modern algebra (105), linear algebra (111), probability and statistics (95), geometry (117), computer science (108), and applications (6). On the other hand, a need to increase certain course offerings was indicated by the following number of schools: analysis (5), modern algebra (10), probability and statistics (28), geometry (38), computer science (50), and applications (47).

Examination of the data presented in Table 2 has led the writers to make some generalizations concerning: (1) how the CUPM recommendations have been met, (2) the increases in course offerings over the past 15 years, and (3) indications of a need to increase course offerings. The results follow:

Calculus — CUPM recommendations are being met. The requirement level has been stable and is expected to remain so.

Analysis — The CUPM recommendations are being met in requirements. The offering level has increased since 1960, but there is little indication that further increases should be anticipated.

Modern Algebra — The CUPM recommendations are being met in requirements. The offering level has increased since 1960, but there is little indication that further increases should be anticipated.

Linear Algebra — Requirements are somewhat short of the CUPM recommendations. There has been an increase in the offering level since 1960, but there is little to no indication that further increases should be anticipated.

Probability and Statistics — Requirements are far short of the CUPM recommendations. There

has been an increase in the offering level since 1960, and there is some indication that further increases should be anticipated.

Geometry — Requirements are only about half of the CUPM recommendations. There has been an increase in the offering level since 1960 and there is some indication that future increases should be anticipated.

Computer Science — Requirements are far short of CUPM recommendations. There has been an increase in the offering level since 1960, but much of this seems to be on the elective rather than the required level. There is strong indication that further increases should be anticipated.

Applications — Requirements are almost non-existent. Few institutions have increased offerings since 1960, but there is a strong indication that increases should be anticipated in the future.

Comments from respondents showed more concern for a course in applications than for any other type of mathematics course. The topics included in such a course would deal with applications of the kind of mathematics which is commonly taught in grades 7–12, as opposed to topics which would be appropriate for engineers, scientists, and the like. The NCTM and the CUPM are presently preparing a mathematics applications sourcebook, but there is a great need for additional applications for mathematics commonly taught in grades 7–12.

The extent to which various organizations are perceived to have influenced changes in preservice programs for secondary school mathematics teachers since 1960 was also investigated. Respondents were asked if they endorsed the Level II-J and Level III recommendations for preservice secondary school mathematics teachers made by the CUPM. Sixty-nine percent of the respondents indicated that they endorsed the CUPM recommendations.

Another part of the questionnaire obtained information as to the extent to which institutions of higher learning in the United States offer courses designed specifically for the preparation of junior high school mathematics teachers. Only 39 respondents indicated that they have at least one course designed specifically for junior high school mathematics teachers. Of those that offered courses designed specifically for junior high school mathematics teachers, only 16 of the respondents indicated that they offered more than one course. In general, courses that were required of senior high school mathematics teachers were also required of junior high school mathematics teachers.

The respondents were asked to list and rank in order of importance the changes that were needed to significantly improve their mathematics content program. Some of these changes and the number of times ranked in regard to importance are given in Table 3.

Thirty-two of the respondents indicated their program's greatest need was for more utilization of the computer in mathematics courses. Twenty-nine said their greatest need was for more geometry, and twenty-six indicated their greatest need was for more applications.

The respondents were also asked to list the mathematics courses that utilized the computer as part of the course. Thirty-six percent of the respondents indicated that the computer was used in numerical analysis, followed by computer science with twenty-nine percent, calculus with 26 percent, statistics with 15 percent, differential equations with ten percent, and linear algebra with nine percent. Thirty-four different courses were listed by the respondents that utilized the computer as part of the course.

Conclusions and recommendations. Institutions of higher learning are offering a larger number of courses in mathematics which allow for more flexibility, creativity, and practicality. In addition, institutions of higher learning are requiring more mathematics content courses, including courses such as computer science, linear algebra, geometry, abstract algebra, and probability and statistics.

Even though only about ten percent of the respondents indicated that their institution of higher learning had at least one course designed specifically for junior high school mathematics teachers, this percentage will increase significantly in the next few years.

The CUPM has been by far the most influential organization since 1960 in regard to changes in preservice secondary school mathematics content programs. It appears that the CUPM Level II-J and

TABLE 3

Changes needed to significantly improve the mathematics content program for preservice secondary school mathematics teachers in 418 institutions of higher learning in the United States in 1974

Needs	Number of Times Ranked				
	1	2	3	4	5
Greater utilization of the computer in mathematics courses	32	10	5	1	1
More geometry	29	5	1	0	0
No needs listed	26	0	0	0	0
More applications	23	5	3	0	0
Course on history of mathematics	10	8	1	0	1
Course on teaching junior high school mathematics	9	3	1	0	0
Better students	9	0	1	0	0
More students	8	2	0	0	0
Course on problem solving	8	0	1	1	0
More required hours for a major in mathematics	7	2	1	0	0
Course on probability and statistics	7	6	1	1	0
More algebra	7	3	0	0	0

Level III recommendations have been fairly well accepted, as indicated by the fact that 69 percent of the respondents stated they endorsed the CUPM recommendations. Further typical programs (based on average requirements plus electives) corresponded fairly well to the CUPM recommendations. The significant changes needed to improve the programs were in directions which, if adopted, would further increase the correspondence between actual programs and CUPM recommendations.

The following recommendations are based upon data obtained for the study and the preceding conclusions:

1. There should be a greater sharing of ideas on how to make more practical applications out of mathematics to real life situations among teachers of mathematics in our institutions of higher learning. This could be done in seminars, professional meetings, and through the professional journals.

2. More mathematics departments in institutions of higher learning should offer courses specifically designed for the preparation of junior high school mathematics teachers.

3. Many schools should offer additional courses in probability and statistics, computer science, geometry, and applications. The applications course should deal more with the applications of secondary school mathematics than does the usual course in applied mathematics. Preservice mathematics teachers should be advised or required to enroll in such courses.

References

1. American Association of Colleges for Teacher Education Directory, The American Association of Colleges for Teacher Education, Washington, D.C., 1973.

2. Christian E. Burckel, The College Blue Book, Christian E. Burckel, New York, 1959.
3. Committee on the Undergraduate Program in Mathematics, Recommendations on Course Content for the Training of Teachers of Mathematics, Mathematical Association of America, 1971.
4. Carl Shepard Johnson, An Analysis of the Required Mathematical Preparation for Secondary School Mathematics Teachers in the United States, Unpublished Doctor's dissertation, Kansas State University, 1975.

Department of Curriculum and Foundations, Marshall University, Huntington, WV 25701.
Department of Curriculum and Instruction, Kansas State University, Manhattan, KS 66506.

MISCELLANEA

MODERN RESEARCH IN MATHEMATICS

A. K. Austin

The advent of Modern Mathematics in the educational limelight has produced an interest in the work of the professional mathematician and in the question, "How can one do research in mathematics?" The following passage formed the introduction to a recent research paper and indicates to some extent the line taken by a number of mathematicians.

A Note on Piffles

A. B. Smith

A. C. Jones in his paper "A Note on the Theory of Boffles", Proceedings of the National Society, 13, first defined a Biffle to be a non-definite Boffle and asked if every Biffle was reducible.

C. D. Brown in "On a paper by A. C. Jones", Biffle, 24, answered in part this question by defining a Wuffle to be a reducible Biffle and he was then able to show that all Wuffles were reducible.

H. Green, P. Smith and D. Jones in their review of Brown's paper, Wuffle Review, 48, suggested the name Woffle for any Wuffle other than the non-trivial Wuffle and conjectured that the total number of Woffles would be at least as great as the number so far known to exist. They asked if this conjecture was the strongest possible.

T. Brown in "A collection of 250 papers on Woffle Theory dedicated to R. S. Green on his 23rd Birthday" defined a Piffle to be an infinite multi-variable sub-polynormal Woffle which does not satisfy the lower regular Q-property. He stated, but was unable to prove, that there were at least a finite number of Piffles.

T. Smith, L. Jones, R. Brown and A. Green in their collected works "A short introduction to the classical theory of the Piffle", Piffle Press, \$20, showed that all bi-universal Piffles were strictly descending and conjectured that to prove a stronger result would be harder.

It is this conjecture which motivated the present paper.

Department of Pure Mathematics, The University, Sheffield, England, S10 2TN

Reprinted, by permission of the author and The Mathematical Association, from The Mathematical Gazette, vol. 51 (1967) pp. 149–150.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before December 31, 1977.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

Errata

E 1243 [1977, 58]. The product $98901 \cdot 10989$ is a perfect square 32967^2, not 21978^2.

E 2440 [1975, 76]. The value of $P(15)$ is incorrect. It should be $P(15) = 74904$.

E 2652 [1977, 295]. The limits for the last summation in the displayed equation should be from i to n, not from 1 to n.

E 2665. *Proposed by Sidney Penner, Bronx Community College, CUNY*

A *partial checkerboard* is a checkerboard from which squares have been removed so that

(a) it is impossible to place even one domino on the remaining board and

(b) the replacement of a single deleted square, regardless of its location, makes it possible to place a domino on the board. (A domino covers two squares having a common side).

It is easy to see that, for an 8×8 partial checkerboard, the minimum number of deleted squares is 32. What is the maximum number?

E 2666. *Proposed by Peter Frankl, Budapest, Hungary*

Let S be a finite set and let $\mathcal{P}$ be the set of all subsets of S. For $\mathcal{A} \subset \mathcal{P}$ and $\mathcal{B} \subset \mathcal{P}$ define $\mathcal{A} * \mathcal{B}$ to be the subset of $\mathcal{P}$ consisting of subsets $X \subset S$ such that $S \subset A \cup B$ for some $A \in \mathcal{A}$ and $B \in \mathcal{B}$.

If $|\mathcal{A}| + |\mathcal{B}| > 2^k$ prove that $|\mathcal{A} * \mathcal{B}| \geqq 2^k$.

E 2667. *Proposed by John R. Samborski, Hyattsville, Maryland*

If $\sum_{k=1}^{\infty} 2^{-n_k}$ is the binary expansion of $(\sqrt{5}-1)/2$ show that $n_k \leqq 5 \cdot 2^{k-2} - 1$.

E 2668. *Proposed by Ron Evans and I. Martin Isaacs, University of Wisconsin.*

Find all non-isosceles triangles with two or more rational sides and with all angles rational (measured in degrees).

E 2669. *Proposed by I. J. Schoenberg, University of Wisconsin*

Let $a > b > 0$. For a given r, $0 < r < b$ there is a unique $R > 0$ such that the circle $(x - a + r)^2 + y^2 = r^2$ lies inside and touches the circle $x^2 + (y - b + R)^2 = R^2$. For which r is R/r minimal?

E 2670*. *Proposed by Shyam Johari, Burroughs Corporation, and Stanley L. Sclove, University of Illinois at Chicago Circle*

Let

$$f(x, y) = \frac{xe^{-x} - ye^{-y}}{e^{-x} - e^{-y}}.$$

If $0 < a < b < c < \infty$ and $0 < x < y < z < \infty$ prove or disprove that

$$|f(a, b) + f(b, c) - f(x, y) - f(y, z)| \leqq 2 \max(|x - a|, |y - b|, |z - c|).$$

SOLUTIONS OF ELEMENTARY PROBLEMS

A Representation for Integers

E 966 [1951, 338; 1976, 378]. *Proposed by C. W. Bruce*

Any integer N may be written in the form $2^k(a^2 - b^2)$, where k, a, b are non-negative integers. When are k, a, b unique? When not unique, how many sets of k, a, b are there?

Solution by H. G. Kopetzky, Leoben, Austria. Let A_N, resp., B_N be the number of representations of N in the form $2^k(a^2 - b^2)$, resp., $a^2 - b^2$ where a, b, k are non-negative integers. We take $N > 0$ (since $A_{-N} = A_N$ and $A_0 = \infty$). Writing $c = a - b$, $d = a + b$ we see that B_N equals the number of factorizations $N = cd$ where $1 \leqq c \leqq d$ and $d - c$ is even. We have $B_N = 0$ if $N \equiv 2 \pmod 4$ and

$$B_N = [\tfrac{1}{2}(\tau(N) + 1)]$$

if N is odd, $\tau(N)$ being the number of positive divisors of N. If $4|N$ then c and d must be even and hence B_N equals the number of factorizations $N/4 = c'd'$ with $1 \leqq c' \leqq d'$, i.e.,

$$B_N = [\tfrac{1}{2}(\tau(N/4) + 1)].$$

If $N = 2^s M$, M odd, then

$$A_N = \sum_{i=0}^{s} B_{2^i M}.$$

Using $\tau(2^i M) = (i + 1)\tau(M)$, a short calculation gives

$$A_N = \tfrac{1}{2}\tau(M)\left(1 + \frac{s(s-1)}{2}\right) + \frac{\varepsilon}{2}\left(1 + \left[\frac{s}{2}\right]\right),$$

where $\varepsilon = 1$ or $\varepsilon = 0$, depending on whether M is a perfect square or not.

It follows that $A_N = 1$ if and only if either $\varepsilon = 0$; $s = 0, 1$; $\tau(M) = 2$ or $\varepsilon = 1$; $s = 0, 1$; $\tau(M) = 1$. That is, $A_N = 1$ only for $N = 1, 2, p$ or $2p$ where p is an odd prime.

Also solved by Karl Beres, D. M. Bloom, Alfred Brousseau, Peter de Buda, Lorraine Foster, S. W. Hahn, William Hoff & Daniel Krider, Elgin Johnston, John Jurgensen & Ray Jurgensen, Ignace Kolodner, Jordan Levy, O. P. Lossers (Netherlands), P. Mallalieu, Catherine Murphy, Marc Passolt, Reinhard Razen (Austria), Eric Rosenthal, Gustavus Simmons, David Stanford, David Stone, G. W. Valk, Kent Wooldridge, and an anonymous solver.

Partially solved by Gary Bates, Javier Erice (Spain), Gustaf Gripenberg (Finland), Eleanor Jones, Margret Kothman, Peter Lindstrom, William Markel, J. W. Mellender, Aaron Meyerowitz (Israel), Victor Pambuccian (Romania), W. F. Smyth (Canada), P. Sriram (India), W. F. Smyth (Canada), and Philip Washburn.

Comment. The above result can be easily deduced from W. Sierpiński, *Elementary Theory of Numbers*, Pańtstrowe Wydawnictwo Naukove, Warszawa 1964, pp. 380–381.

A Locus Associated with Two Segments

E 1822 [1965, 903; 1976, 53]. *Proposed by Necdet Ucoluk*

Let A, A_1 and B, B_1 be any two pairs of given points in the plane. Consider the locus of points N such that the angles ANA_1 and BNB_1 (with measures having absolute values α and β respectively) satisfy the equation $\alpha = k\beta$, where k is a given positive real number.

(a) Determine the differentiability properties of this locus, and

(b) when the tangent line exists give a geometric procedure (finite) for its construction.

Solution by Anthony G. O'Farrell, St. Patrick's College, Kildare, Ireland. Let a, a_1, b, b_1, z be the complex numbers representing the points A, A_1, B, B_1, N, respectively. The above locus Γ is then locally expressible as

$$\arg(z-a)-\arg(z-a_1) = \pm k\,(\arg(z-b)-\arg(z-b_1)) \tag{1}$$

with appropriate determination of the arguments involved.

First consider the part Γ_+ of Γ given by the $+$ sign. For $z \neq a$, a_1, b, b_1 let

$$f(z) = \arg(z-a)-\arg(z-a_1)-k\,(\arg(z-b)-\arg(z-b_1)).$$

Since

$$\frac{\partial}{\partial x}\arg(z-a) = \frac{\partial}{\partial x}\operatorname{Im}\log(z-a) = \operatorname{Im}\frac{\partial}{\partial x}\log(z-a) = \operatorname{Im}\left(\frac{1}{z-a}\right),$$

$$\frac{\partial}{\partial y}\arg(z-a) = -\operatorname{Re}\left(\frac{1}{z-a}\right),$$

the gradient of $\arg(z-a)$ is given by

$$\nabla\arg(z-a) = i\cdot\frac{1}{\bar{z}-\bar{a}}.$$

Hence $(\nabla f)(z) = i\overline{\theta(z)}$ where

$$\theta(z) = \frac{1}{z-a}-\frac{1}{z-a_1}-k\left(\frac{1}{z-b}-\frac{1}{z-b_1}\right).$$

If $z_0 \in \Gamma_+$ and $\theta(z_0) \neq 0$ then Γ_+ has a tangent at z_0 and its direction is given by $\overline{\theta(z_0)}$. It is clear that $\theta(z_0)$ can be easily constructed when a, a_1, b, b_1 and z_0 are known.

The critical points z are those points of Γ_+ for which $\theta(z) = 0$, i.e.,

$$\frac{(z-b)(z-b_1)}{(z-a)(z-a_1)} = k\frac{b-b_1}{a-a_1}.$$

The part Γ_- of Γ corresponding to the $-$ sign in (1) can be treated similarly.

Also solved by Jordi Dou (Spain).

Comments. Dou remarks that Γ has cusps at a, a_1, b, b_1; that Γ_+ and Γ_- intersect at the point where the lines AA_1 and BB_1 meet and that Γ has two asymptotes. Let

$$c = \frac{a+kb}{1+k},\ c_1 = \frac{a_1+kb_1}{1+k},\ d = \frac{a+kb_1}{1+k},\ d_1 = \frac{a_1+kb}{1+k}.$$

The two asymptotes are the lines through c, c_1 and d, d_1.

A Periodic Recurrence

E 2567 [1975, 1010]. *Proposed by J. H. Conway, Cambridge University, England and R. L. Graham, Bell Laboratories, Murray Hill, New Jersey*

Define polynomials $f_m = f_m(x_1,\ldots,x_m)$ by $f_0 = 1$, $f_1 = x_1$, $f_k = x_k f_{k-1} - f_{k-2}$, $k \geqq 2$. For a fixed $n \geqq 3$, let $y_1, y_2, \ldots$ satisfy $f_n(y_{k+1},\ldots,y_{k+n}) = 1$ for all $k \geqq 0$. Show that $y_{n+k+2} = y_k$ for all $k \geqq 1$.

Solution by Gérard Letac, Université Paul Sabatier, Toulouse, France (revised by the Editor). Let $x = (x_1, x_2, \ldots)$, $y = (y_1, y_2, \ldots)$. For any infinite sequence $a = (a_1, a_2, \ldots)$ we shall write $Sa = (a_2, a_3, \ldots)$. We define $f_m(x) = f_m = f_m(x_1, x_2, \ldots, x_m)$ and similarly $f_m(y)$. By hypothesis we have

$$f_n(S^k y) = 1 \qquad (k \geqq 0). \tag{1}$$

We shall prove the above assertion under the additional hypotheses

$$f_{n-1}(S^k y) \neq 0 \qquad (1 \leqq k \leqq n+2). \tag{2}$$

Define

$$M_k(x) = \begin{pmatrix} x_k & -1 \\ 1 & 0 \end{pmatrix}.$$

It is easy to prove by induction that

$$M_p(x) M_{p-1}(x) \cdots M_1(x) = \begin{pmatrix} f_p(x) & -f_{p-1}(Sx) \\ f_{p-1}(x) & -f_{p-2}(Sx) \end{pmatrix}. \tag{3}$$

Replacing x by $S^k y$ and taking determinants on both sides, we obtain

$$f_{p-1}(S^k y) f_{p-1}(S^{k+1} y) - f_p(S^k y) f_{p-2}(S^{k+1} y) = 1. \tag{4}$$

Taking $p = n+1$ in (4) and using (1) and (2), we obtain

$$f_{n+1}(S^k y) = 0, \qquad 0 \leqq k \leqq n+1. \tag{5}$$

Taking $p = n+2$ in (4) and using (1) and (5), we obtain

$$f_{n+2}(S^k y) = -1, \qquad 0 \leqq k \leqq n+1. \tag{6}$$

Replacing x by $S^k y$ in (3) and taking $p = n+2$, we obtain

$$M_{n+k+2}(y) M_{n+k+1}(y) \cdots M_{k+1}(y) = -I, \qquad 0 \leqq k \leqq n$$

because of (5), (6) and since $M_m(S^k y) = M_{m+k}(y)$. These matrix equalities imply that $M_{n+k+2}(y) = M_k(y)$, i.e., $y_{n+k+2} = y_k$ for $1 \leqq k \leqq n$.

Using this it is easy to check that the sequence $z = Sy$ also satisfies the conditions (1) and (2). Hence, we can conclude that also $z_n = z_{2n+2}$, $y_{n+1} = y_{2n+3}$. This argument can be repeated ad infinitum to get $y_{n+k+2} = y_k$ for all $k \geqq 1$.

Also solved by David Cantor, L. E. Mattics, J. G. Mauldon, Bryce Parry, Michael Skalsky, J. G. Wendel, and the proposers.

Editor's Comment. All solvers provided counterexamples in case where no additional conditions, such as (2), are imposed. Reinhard Piater and Martin Schechter have also given such counterexamples.

Characterizing Solutions of a Functional Equation

E 2583 [1976, 198]. *Proposed by C. L. Mallows, Bell Laboratories, Murray Hill, New Jersey*

Find all continuous functions $g:\mathbf{R}\to\mathbf{R}$ such that, for some continuous $f:\mathbf{R}^2\to\mathbf{R}$, we have $g(xy)=f(x,g(y))$ for all x, y in $\mathbf{R}$.

Solution by Peter L. Montgomery, Huntsville, Alabama. We claim that this is the case if and only if g satisfies one of the following:

(i) g is constant;

(ii) g is bijective;

(iii) g is even and its restriction g_1 to $[0,+\infty)$ is strictly monotonic and unbounded.

Sufficiency. In case (i) we take $f(x,y)=y$. In case (ii) we take $f(x,y)=g(xg^{-1}(y))$. In case (iii) we define $h:\mathbf{R}\to\mathbf{R}$ by

$$h(x)=\begin{cases} g_1^{-1}(x) & \text{for } x\in g(\mathbf{R})=g_1(\mathbf{R})\\ 0 & \text{otherwise,}\end{cases}$$

and we can take $f(x,y)=g(xh(y))$.

Necessity. Assume first that g is injective. Since g is continuous, it follows that g is strictly monotonic. We claim that in fact g is bijective. For this it suffices to show that $a=\lim_{x\to+\infty}g(x)$ and $b=\lim_{x\to-\infty}g(x)$ are not finite. If a were finite, then

$$g(x)=f(xy^{-1},g(y))$$

shows that $g(x)=f(0,a)$ for all x. This is a contradiction. Similar contradiction is obtained if b is finite. Thus (ii) holds.

Now assume that g is neither injective nor constant. If $g(x)=g(y)$ and $|x|\leqq|y|\neq 0$, then for $t=xy^{-1}$ and $z\in\mathbf{R}$ we have

$$g(tz)=g(xy^{-1}z)=f(y^{-1}z,g(x))=f(y^{-1}z,g(y))=g(z).$$

Hence $g(z)=g(t^nz)$ for $n=1,2,3,\ldots$. If $|x|<|y|$ then $|t|<1$ and since g is continuous at 0 we obtain the contradiction $g=\text{constant}$. Hence $g(x)=g(y)$ implies $y=\pm x$. Since g is not injective, there exists $x>0$ such that $g(-x)=g(x)$. Thus, taking $y=-x$ we have $t=-1$ and $g(-z)=g(z)$ for all $z\in\mathbf{R}$, i.e., g is even. Moreover, the restriction g_1 of g to $[0,+\infty)$ is injective. Thus g_1 is strictly monotonic and if $a=\lim_{x\to+\infty}g_1(x)$ then a cannot be finite (the proof is the same as in the case when g is injective).

Also solved by Mangho Ahuja, Dwight Bean, Irl Bivens, John Bryant & Robert Gilmer, Peter de Buda (Canada), James Gard, G. A. Heuer, Joel Levy, O. P. Lossers (Netherlands), L. E. Mattics, David Stanford, J. B. Wilker (England), and the proposer.

A Matrix Squared

E 2586 [1976, 198]. *Proposed by Walter Egerland, U.S. Army Ballistic Research Laboratories, Aberdeen, Maryland*

Evaluate $\det A$ where $A=(a_{ij})$ is the $(n+1)\times(n+1)$ matrix defined by

$$
\begin{aligned}
&a_{ij} = 0 && \text{if } i - j \neq 0, \pm 2,\\
&a_{ii} = \lambda_i + \lambda_{i-1} && (\text{where } \lambda_0 = \lambda_{n+1} = 0),\\
&a_{i+2,i} = 1, && a_{i,i+2} = \lambda_i\lambda_{i+1}
\end{aligned}
$$

(The scalars λ_i may belong to any commutative ring.)

Solution by Angel Santos Paloma, Malaga, Spain. Note that $A = B_{n+1}^2$ where B_{n+1} is the $(n+1)\times(n+1)$ matrix

$$
\begin{pmatrix}
0 & \lambda_1 & 0 \cdots & & & \\
1 & 0 & \lambda_2 \cdots & 0 & & \\
0 & 1 & 0 \cdots & & & \\
\vdots & \vdots & \vdots & 0 & \lambda_{n-1} & 0\\
 & & & 1 & 0 & \lambda_n\\
 & 0 & & 0 & 1 & 0
\end{pmatrix}
$$

Expanding by the last column, one finds that

$$\det B_{n+1} = -\lambda_n \det B_{n-1}.$$

Since $\det B_1 = 0$, $\det B_2 = -\lambda_1$, the above recursive formula implies that

$$
\det B_{n+1} = \begin{cases} 0 & \text{if } n \text{ is even}\\ (-1)^{(n+1)/2}\lambda_n\lambda_{n-2}\cdots\lambda_3\lambda_1 & \text{if } n \text{ is odd.}\end{cases}
$$

Also solved by Marcia Ascher, The Bennett College Team, Paul Chauveheid (Belgium), José Luis de Miguel (Spain), Thomas Foregger, Ivey Gentry, Clark Givens, M. G. Greening (Australia), Anita Grossman, Frederick Humburg, A. A. Jagers (Netherlands), J. D. Jones, L. Kuipers (Switzerland), David Lantz, Carolyn MacDonald, L. E. Mattics, Rajan Modi, Peter Montgomery, Robert Patenaude, A. J. Roques, Daniel Rosenblum, Chander Sabharwal & Chaman Sabharwal, August Sardinas, University of South Alabama Problem Group, and the proposer.

An Application of Brouwer's Fixed Point Theorem

E 2587 [1976, 284]. *Proposed by Bruno O. Shubert, Naval Postgraduate School*

Consider the system of n equations

$$x_0 + x_k = \min_{j=1,\ldots,m}\ \max_{i=1,\ldots,n}\ (a_{ijk} + x_i), \qquad k = 1,\ldots,n,$$

in $n+1$ unknowns $x_0, x_1, \ldots, x_n$, where the a_{ijk} are given constants. Show that

(1) the system always has a solution, and

(2) the first component x_0 is unique.

Solution by the proposer. Let $a = \max_{i,j,k} |a_{ijk}|$;

$$A = \{x \in \mathbf{R}^n \mid \sum_{i=1}^{n} x_i = 0, \qquad \max |x_i| \leqq 2a\};$$

$$f_k(x)=\min_j \max_i (a_{ijk}+x_i), \qquad 1\leqq k\leqq n;$$

$$g_k(x)=f_k(x)-\frac{1}{n}\sum_{r=1}^{n} f_r(x).$$

For $x\in\mathbf{R}^n$ we have

$$\begin{aligned}|f_k(x)-f_r(x)| &= \left|\min_j \max_i (a_{ijk}+x_i)-\min_j \max_i (a_{ijr}+x_i)\right|\\ &\leqq \max_j \left|\max_i (a_{ijk}+x_i)-\max_i (a_{ijr}+x_i)\right|\\ &\leqq \max_j \max_i |a_{ijk}-a_{ijr}|\leqq 2a,\end{aligned}$$

and

$$|g_k(x)|\leqq \frac{1}{n}\sum_{r=1}^{n}|f_k(x)-f_r(x)|\leqq 2a.$$

Thus, $g=(g_1,\dots,g_n)$ maps A into itself and by Brouwer's fixed point theorem there is a $b\in A$ such that $g(b)=b$. By taking

$$b_0=\frac{1}{n}\sum_{i=1}^{n} f_i(b)$$

we obtain a solution $(b_0,b_1,\dots,b_n)$ of the given system.

Let $(c_0,c_1,\dots,c_n)$ be also a solution. Note that $h_k(x)=f_k(x)-x_k$ is an increasing function of x_i for $i\neq k$. Hence, if c_i-b_i $(1\leqq i\leqq n)$ takes its maximal value $c_s-b_s=\delta$ for $i=s$, we have

$$\begin{aligned}c_0-b_0&=h_s(c_1,\dots,c_n)-h_s(b_1,\dots,b_n)=h_s(c_1-\delta,\dots,c_n-\delta)-h_s(b_1,\dots,b_n)\\ &\leqq h_s(b_1,\dots,b_{s-1},c_s-\delta,b_{s+1},\dots,b_n)-h_s(b_1,\dots,b_n)=0\end{aligned}$$

since $c_i-\delta\leqq b_i$ for $i\neq s$ and $c_s-\delta=b_s$. Thus $c_0\leqq b_0$ and similarly $b_0\leqq c_0$, so that $b_0=c_0$.

Computation of a Determinant

E 2588 [1976, 284]. *Proposed by Stephen B. Maurer, Princeton University*

Let A_n be the matrix of order $(2^n-1)\times n$ whose kth row is the binary expression for k. Let $M_n=A_nA_n^t \pmod 2$. If M_n is regarded as a matrix over the integers, what is its determinant?

Solution by Jerry Griggs and Bruce Sagan, Massachusetts Institute of Technology Combinatorics Class. We have $\det M_1=1$. For $n\geqq 1$, by considering how M_{n+1} is constructed, we have:

$$M_{n+1}=\begin{pmatrix} & & 0 & & \\ & M_n & \vdots & M_n & \\ & & 0 & & \\ 0 & 0\cdots 0 & 1 & 1\cdots 1 & 1\\ & & 1 & & \\ & M_n & \vdots & J-M_n & \\ & & 1 & & \end{pmatrix}$$

where J is the square matrix of all ones of dimensions $2^n - 1$. Subtract the middle column from all columns to its right. Then

$$\det M_{n+1} = \begin{vmatrix} & 0 & \\ & 0 & \\ M_n & \vdots & M_n \\ & 0 & \\ 0\cdots 0\cdots 0 & 1 & 0\cdots 0\cdots 0 \\ M_n & \vdots & -M_n \\ & 1 & \end{vmatrix} = \begin{vmatrix} M_n & M_n \\ M_n & -M_n \end{vmatrix}$$

$$= \begin{vmatrix} 2M_n & 0 \\ M_n & -M_n \end{vmatrix} = (-2)^{2^n-1}(\det M_n)^2.$$

Solving this simple recursion gives the result we seek

$$\det M_n = -2^{(n-2)2^{n-1}+1} \qquad (n > 1).$$

Also solved by Robert Brigham, Peter de Buda, José Luis de Miguel (Spain), Thomas Foregger, S. D. Godse & K. R. P. Singh (India), A. A. Jagers (Netherlands), J. D. Jones, O. P. Lossers (Netherlands), Thomas McCormick, Aaron Meyerowitz (Israel), Rajan Modi, David Reiner, Eric Rosenthal, E. G. Strauss, University of South Alabama Problem Group, Gillian Valk, and the proposer. Partial solutions by Peter Liepa (Canada), and Unni Namboodiri.

Another Determinant

E 2589 [1976, 284]. *Proposed by Joe Sunday, University of Guelph, Ontario*

Let $d_1, \ldots, d_n$ be distinct integers > 1. If $a_{ij} = \sin^2(j\pi/d_i)$ for $1 \leqq i, j \leqq n$, show that $\det(a_{ij}) \neq 0$.

Solution by J. B. Wilker, University of Exeter, England. We shall compute the determinant of the matrix (a_{ij}), $1 \leqq i, j \leqq n$, where $a_{ij} = \sin^2(j\theta_i)$ and $\theta_1, \ldots, \theta_n$ are real numbers.

Since $\sin^2 j\theta$ is a polynomial in $\sin^2\theta$ with leading coefficient $(-4)^{j-1}$ we have

$$\det(a_{ij}) = (-4)^{n(n-1)/2}\det(b_{ij})$$

with $b_{ij} = \sin^{2j}\theta_i$. By reversing the order of rows of (b_{ij}) we obtain

$$\det(a_{ij}) = 2^{n(n-1)}\det(c_{ij})$$

where $c_{ij} = b_{n-i+1,j}$. Using the known expression for the determinant of the Vandermonde matrix we find that

$$\det(a_{ij}) = 2^{n(n-1)}\prod_{i=1}^{n}\sin^2\theta_i \prod_{i<j}(\sin^2\theta_i - \sin^2\theta_j).$$

Hence if $\theta_1, \ldots, \theta_n$ are distinct and satisfy $0 < \theta_i \leqq \pi/2$ then (a_{ij}) is non-singular.

Also solved by H. A. al-Tayyar & Y. A. Said (Iraq), Robert Breusch, Ron Evans, Emden Gansner, A. A. Jagers (Netherlands), L. E. Mattics, and St. Olaf Problem Group.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Hill Center, Rutgers University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate signed sheets and should be mailed before December 31, 1977.
An asterisk (*) means neither the proposer nor the editors supplied a solution.

6162. *Proposed by the late Ray Latham*

If $A = (a_{ij})$ is the $n \times n$ matrix defined by $a_{ij} = 1/(1-4(i-j)^2)$, and $\mathbf{x} = (x_i)$ is the unique vector such that $A\mathbf{x} = \mathbf{e}$ (the all 1's vector), show that

$$\sum_{i=1}^{n} x_i = \binom{n+1}{2}.$$

6163. *Proposed by John Myhill, State University of New York at Buffalo*

Devise an algorithm for escaping from a connected, countably infinite, locally finite maze. "Countably infinite" means the number of edges and nodes is $\aleph_0$, "locally finite" means only a finite number of edges meet at each node. "Algorithm" means this: You are lost in the middle of the maze, having no idea of where the exit is. Your only possibility of escape therefore is to devise a tour which will take you through every node of the maze after a finite number of steps. In order to keep track of your route, you are given an everlasting pencil and an infallible eraser; at each node, and at the roadside of each road near the node, is a board on which you can write and erase. However, you have only a finite alphabet to write with, and there is a fixed bound on how many characters you can write on the boards. (In particular, then, you cannot keep on any board a record of how many times you have passed it.) Your field of vision is limited to being able to see, from any node, what is written on the board at that node and what is written on the nearby roadside boards.

(*) Can this procedure be altered to solve the locally infinite case?

6164. *Proposed by Ignacy I. Kotlarski, Oklahoma State University*

Let the random variable $Z_1 = X$ follow the Cauchy distribution with the probability density function $f(x) = [\pi(1+x^2)]^{-1}$, $x \in R$. Show that for $n = 2, 3, \ldots$, the random variables

$$Z_2 = \frac{2X}{1-X^2}, \quad Z_3 = \frac{3X - X^3}{1-3X^2}, \quad Z_4 = \frac{4X-4X^3}{1-6X^2+X^4}, \cdots$$

$$Z_n = \frac{\binom{n}{1}X - \binom{n}{3}X^3 + \binom{n}{5}X^5 - \cdots}{1 - \binom{n}{2}X^2 + \binom{n}{4}X^4 - \binom{n}{6}X^6 + \cdots}, \ldots$$

also follow the same Cauchy distribution.

6165. *Proposed by A. G. O'Farrell, St. Patrick's College, Kildare, Ireland*

Suppose $f(x)$ is a real-valued function on $\mathbb{R}^n$ and define

$$M(x,r) = \frac{\int_{|x-y|\leq r} f(y)\,dy}{\int_{|x-y|\leq r} 1\,dy}, \quad \text{for} \quad x \in \mathbb{R}^n, \qquad r > 0.$$

Suppose

$$\frac{M(x,r) - f(x)}{r^2} \to 0 \quad \text{as} \quad r \downarrow 0$$

for each $x \in \mathbb{R}^n$. Show that $f(x)$ is harmonic.

6166. *Proposed by D. A. Gregory, Queen's University, Kingston, Ontario*

If f is a convex functional on a convex subset K of a vector space, then for all x and $x+h$ in K, the one-sided directional derivatives

$$f'_+(x,h)=\lim_{\alpha\to 0^+}\frac{f(x+\alpha h)-f(x)}{\alpha}$$

exist in the extended reals and $f(x+h)\geqq f(x)+f'_+(x,h)$. Is the converse true? If so, we have an analytic characterization of convex functionals.

6167.* *Proposed by Charles R. Williams and Joseph C. Warndof, Midwestern University*

Suppose $f\colon R^n \to R^{n-1}$ and for each point $a \in R^n$, the limit

$$\lim_{x\to a}\frac{|f(x)-f(a)|}{|x-a|}$$

exists. Is f necessarily a constant function? (The answer is "yes" if $n=2$.)

SOLUTIONS OF ADVANCED PROBLEMS

Consecutive Quadratic Nonresidues

6058 [1975, 942]. *Proposed by Larry Taylor, New York, N.Y.*

(A) If $p\equiv 31$ or $39 \pmod{40}$ is prime and if $a\equiv(\sqrt{5}+2)/3$ and $b\equiv(\sqrt{5}-2)/3$ are of even order $(\operatorname{mod} p)$, prove that either $a-1$, a and $a+1$ or $b-1$, b and $b+1$ are quadratic nonresidues of p.

(B) If $p\equiv 19 \pmod{24}$ is prime and $a\equiv\sqrt{-1/3}$ is of even order $(\operatorname{mod} p)$, prove that $a-1$, a and $a+1$ are quadratic nonresidues of p.

(For example, (A) 21, 22, and 23 are quadratic nonresidues of 31; (B) 32, 33 and 34 are quadratic nonresidues of 43.)

Solution by Lorraine L. Foster, California State University at Northridge. All symbols (x/p) in the following are Legendre symbols.

(A) Let $p\equiv 31$ or $39 \pmod{40}$. Then $(-1/p)=-1$, $(2/p)=1$, $(5/p)=1$. Since $(p-1)/2$ is odd and $x^{(p-1)/2}\equiv(x/p) \pmod p$, x of even order implies $(x/p)=-1$. Suppose $a\equiv(\sqrt{5}+2)/3$ and $b\equiv(\sqrt{5}-2)/3 \pmod p$ are of even order so that $(a/p)=(b/p)=-1$. Now

$$\left(\frac{(a-1)(b+1)}{p}\right)=\left(\frac{4/9}{p}\right)=1 \text{ so that } \left(\frac{a-1}{p}\right)=\left(\frac{b+1}{p}\right)=\pm 1.$$

Also

$$\left(\frac{(a+1)(b-1)}{p}\right)=\left(\frac{-5(4)/9}{p}\right)=-1$$

so that $((a+1)/p)=-((b-1)/p)\neq 0$. Further, $9(b+1)^2 b\equiv 2(a-1) \pmod p$ so that

$$\left(\frac{9(b+1)^2}{p}\right)\left(\frac{b}{p}\right)=\left(\frac{2}{p}\right)\left(\frac{a-1}{p}\right),\quad \left(\frac{a-1}{p}\right)=-1.$$

Thus, since $(a/p)=(b/p)=((a-1)/p)=((b+1)/p)=-1$ and since $((a+1)/p)=-((b-1)/p)$, exactly one of the sets $\{a-1,a,a+1\}$, $\{b-1,b,b+1\}$ consists of all quadratic nonresidues.

(B) Let $p=24k+19$. Then

$$\left(\frac{-1}{p}\right)=-1,\quad \left(\frac{2}{p}\right)=-1,\quad \left(\frac{3}{p}\right)=\left(\frac{1/3}{p}\right)=-1,\quad \left(\frac{-1/3}{p}\right)=1.$$

Let $a\equiv\sqrt{-1/3}\,(\operatorname{mod} p)$ be of even order $\operatorname{mod} p$. Then, as above, $(a/p)=-1$. Now

$$\left(\frac{a+1}{p}\right)\left(\frac{a-1}{p}\right)=\left(\frac{-4/3}{p}\right)=1 \text{ so that } \left(\frac{a+1}{p}\right)=\left(\frac{a-1}{p}\right)=\pm 1.$$

Also, $(a+1)^2\equiv -2a(a-1)\,(\operatorname{mod} p)$ so that

$$1=\left(\frac{(a+1)^2}{p}\right)=\left(\frac{-1}{p}\right)\left(\frac{2}{p}\right)\left(\frac{a}{p}\right)\left(\frac{a-1}{p}\right)=(-1)^3\left(\frac{a-1}{p}\right)$$

or $((a-1)/p)=-1$, $((a+1)/p)=-1$.

Also solved by Paul Bruckman, M. G. Greening (Australia), L. Kuipers (Switzerland), A. A. Jagers (Netherlands), L. E. Mattics, and the proposer.

Note. For additional primitive roots see Taylor and Shanks, *An Observation of Fibonacci primitive roots,* Fibonacci Quarterly, vii, #2 (1973).

Cyclic Sylow Subgroups of Metacyclic Groups

6059 [1975, 942]. *Proposed by S. Baskaran, University of Madras, India*

In his book, *The Theory of Groups,* M. Hall calls a group G metacylic if the derived group G' and the factor group G/G' are both cyclic. Prove that if G is a finite metacylic group and p is the smallest prime dividing the order of G then a Sylow p-subgroup of G is cyclic.

Solution by Roger D. Peterson and Jay I. Miller, University of Wisconsin at Milwaukee. G is supersolvable because it is metacylic. Hence G has a normal p-complement, N. Then $G=NP$ with $N\lhd G$ and $N\cap P=1$, where P is a Sylow p-subgroup of G. But then $G/N\cong P$ and $P/P'\cong(G/N)/(G/N)'\cong(G/N)/((G'N)/N)\cong G/G'N$, so that P/P' is cyclic. An easy argument shows that P is cyclic (See B. Huppert, *Endliche Gruppen I,* Springer-Verlag, 1967, III, 7.1).

Note. A more general definition is that G is metacylic provided G has a normal subgroup H such that both H and G/H are cyclic. With this definition, the above result fails, since there are noncyclic metacylic p-groups for every prime p.

Also solved by Henry Bray, David Buchthal, A. A. Jagers (Netherlands), Hans Liebeck (England), Harvey Schmidt, Jr., Armond Spencer, R. M. Stafford & F. X. Felcon, and the proposer.

Note. Other solvers observe that the condition that p is the smallest prime dividing $|G|$ is necessary, since otherwise $S_3\times Z_3$ would be a counterexample.

A Convex Collection of $n\times n$ Matrices

6061 [1975, 1016]. *Proposed by Hung C. Li, Southern Colorado State College*

For any positive semi-definite Hermitian matrix H, $(n\times n)$, the set

$$S=\{A\,|\operatorname{tr}(AA^*)H\leqq\lambda\}$$

is convex in A, where A is $n\times m$, X^* is the complex conjugate and transpose of X, and $\operatorname{tr}X$ is the trace of X.

Solution by Ingram Olkin, Stanford University. We prove a stronger result, namely, that $G(A)=AA^*$ is a convex function of A in the sense that

$$(*)\qquad G(\alpha A+\bar{\alpha}B)\leqq\alpha G(A)+\bar{\alpha}G(B)\quad\text{for}\quad 0\leqq\alpha\leqq 1,\quad \alpha+\bar{\alpha}=1,$$

where $U\leqq V$ means $V-U$ is positive semi-definite (p.s.d.)

We show directly that

$$(1)\qquad (\alpha A+\bar{\alpha}B)(\alpha A+\bar{\alpha}B)^* \leqq \alpha AA^*+\bar{\alpha}BB^*,$$

for this is equivalent to

$$\alpha^2 AA^*+\alpha\bar{\alpha}(BA^*+AB^*)+\bar{\alpha}^2 BB^* \leqq \alpha AA^*+\bar{\alpha}BB^*.$$

Upon simplification, this is

$$0\leqq \alpha\bar{\alpha}(AA^*+BB^*-BA^*-AB^*)=\alpha\bar{\alpha}(A-B)(A-B)^*,$$

which clearly holds.

As a consequence of (*) it follows that for any p.s.d. Hermitian matrix H with p.s.d. Hermitian square root $H^{1/2}$,

$$(2)\qquad H^{1/2}G(\alpha A+\bar{\alpha}B)H^{1/2}\leqq \alpha H^{1/2}G(A)H^{1/2}+\bar{\alpha}H^{1/2}G(B)H^{1/2}.$$

Since the trace is linear, (2) implies that

$$\operatorname{tr} G(\alpha A+\bar{\alpha}B)H\leqq \alpha \operatorname{tr} G(A)H+\bar{\alpha}\operatorname{tr} G(B)H,$$

which in turn implies the result.

Also solved by A. J. Bosch (Netherlands), Chandler Davis, Clyde Davis, Emeric Deutsch, Thomas Foregger, Asako Higa (Japan), Marvin Marcus, A. McD. Mercer (Canada), William Watkins, and the proposer.

Nesting Regular n-gons

6062* [1975, 1016]. *Proposed by B. H. Voorhees, University of Alberta, Canada*

Consider an infinite sequence of regular n-gons such that each $(n+1)$-gon is contained within the preceding n-gon and is of maximal area consistent with this constraint. Take the first element of this sequence as an equilateral triangle having unit area. Is the limit of this sequence a point or a circle? If it is a circle, determine its area.

Partial Solution by E. F. Schmeichel and J. G. Pierce, University of Southern California. The limit of the given sequence is not a point. To prove this, it suffices to observe that the following sequence of regular figures converges to a circle: an equilateral triangle, a circle inscribed in this triangle, a square inscribed in this circle, a circle inscribed in the square, etc.

Let r_n denote the radius of the circle inscribed in the regular $(n-1)$-gon. It is easily seen that $r_{n+1}/r_n=\cos(\pi/n)$, for $n\geqq 4$, and so $r_n=r_4\prod_{k=4}^{n}\cos(\pi/k)$. Hence

$$\lim_{n\to\infty} r_n=r_4\prod_{k=4}^{\infty}\cos(\pi/k)>r_4\prod_{k=4}^{\infty}(1-\tfrac{1}{2}(\pi/k)^2)>0.$$

See Bromwich, *Infinite Series*, Art. 39). Thus this particular sequence of figures converges to a circle.

Also solved (partially) by Bo Beradkson (Netherlands), Paul Erdös (Israel), Lawrence House, and Joyce Williams.

Editorial Note. It will be noted that inscribing an $(n+1)$-gon in a circle inscribed in the regular n-gon does not give the largest $(n+1)$-gon inside the n-gon. (Consider the square in the equilateral triangle.) No indication as to the desired limiting area has been contributed.

Distance Between the Centers of Two Spheres

6063 [1975, 1016]. *Proposed by H. J. Marcum, Universidade Federal do Rio de Janeiro, Brazil*

Let S be the set of all circles in the plane provided with the Hausdorff metric ρ induced by the

usual Euclidean metric d, (i.e., $\rho(A,B)=\max\{\sup_{a\in A} d(a,B),\ \sup_{b\in B} d(b,A)\}$, where $d(a,B)=\inf\{d(a,b): b\in B\}$ denotes the distance from the point a to the set B).

Let $z: S\to\mathbb{R}^2$ be the function which to each circle A assigns its center $z(A)$. Prove that $d(z(A),z(B))\leqq\rho(A,B)$ for all $A,B\in S$.

I. Solution by the University of South Alabama Problem Group. Let r_1 be the radius of circle A and r_2 be the radius of circle B and let $s=d(z(A),z(B))$. If one looks at the six possible cases (such as A and B intersect and $z(A)$ is within B but $z(B)$ is not within A) one finds that

$$\rho(A,B)=\max\{s+r_1-r_2,\ s+r_2-r_1\}\geqq s.$$

II. Solution by the proposer. Consider A,B as geometric k-spheres in $\mathbb{R}^n$. Recall the given definition of $\rho(A,B)$ and suppose that $z(A)\neq z(B)$. Let L_A be the $(n-1)$-plane through $z(A)$ perpendicular to $z(B)$ and let L_B be the $(n-1)$-plane through $z(B)$ perpendicular to $z(A)$. Put

$$\Delta(A,L_A)=\sup\{d(a,L_A): a\in A\},\quad \Delta(B,L_B)=\sup\{d(b,L_B): b\in B\}.$$

Note first that for any $x\in\mathbb{R}^n$, $d(x,L_A)-\Delta(A,L_A)\leqq d(x,A)$ since

$$\begin{aligned}d(x,L_A)&\leqq d(x,a_0)+d(a_0,L_A)=d(x,A)+d(a_0,L_A)\\&\leqq d(x,A)+\sup\{d(a,L_A): a\in A\}=d(x,A)+\Delta(A,L_A)\end{aligned}$$

where $a_0\in A$ is such that $d(x,A)=d(x,a_0)$. Similarly $d(x,L_B)-\Delta(B,L_B)\leqq d(x,B)$ for any $x\in\mathbb{R}^n$.

Now on the line passing through $z(A)$ and $z(B)$ let e_A and e_B be the unique points such that $z(A)$ lies between e_A and $z(B)$, and $z(B)$ lies between e_B and $z(A)$ and such that $d(z(A),e_A)=$ the radius of A, $d(z(B),e_B)=$ the radius of B. Next choose $a_1\in A$ and $b_1\in B$ such that $d(e_A,A)=d(e_A,a_1)$ and $d(e_B,B)=d(e_B,b_1)$. We consider two cases. If $\Delta(A,L_A)\leqq\Delta(B,L_B)$ then $d(b_1,L_A)=d(z(A),z(B))+d(b_1,L_B)$ and $d(b_1,L_B)=\Delta(B,L_B)$, so that

$$\begin{aligned}d(z(A),z(B))&=d(b_1,L_A)-d(b_1,L_B)=d(b_1,L_A)-\Delta(B,L_B)\\&\leqq d(b_1,L_A)-\Delta(A,L_A)\leqq d(b_1,A)\\&\leqq\sup\{d(b,A): b\in B\}\leqq\rho(A,B).\end{aligned}$$

On the other hand, if $\Delta(B,L_B)\leqq\Delta(A,L_A)$ then we may argue as in the previous case by interchanging A and B to show that $d(z(A),z(B))\leqq\rho(A,B)$ and so complete the proof.

Also solved by R. A. Christiansen, Emeric Deutsch, Joel Levy, O. P. Lossers (Netherlands), Alvin Martin, James Munkres, Randy Schilling, Kenneth Smith, Arthur Solomon, R. H. Sorgenfrey, Rick Troxel and R. M. Warten.

Editor's Notes. (1) It had been the proposer's intent that the problem be posed for the more general case in II above. (2) Munkres considered a more general case by varying the metric: Let d be a metric for $\mathbb{R}^n$ satisfying the conditions

(1) $d(a,b)+d(b,c)=d(a,c)$ if b lies on the line segment ac.

(2) d is unbounded on each ray.

[Examples of such metrics include any metric derived from a norm on $\mathbb{R}^n$ such as

$$d(x,y)=(\Sigma|x_i-y_i|^p)^{1/p}\quad\text{for } p\geqq 1,\quad\text{and}$$

$$d(x,y)=\max\{|x_i-y_i|\}.]$$

Given $z\in\mathbb{R}^n$ and given $r>0$, we call the set $A(z,r)=\{x\mid d(x,z)=r\}$ the (generalized) *circle* with center z and radius r. Let S be the set of all (generalized) circles in $\mathbb{R}^n$, provided with the Hausdorff metric ρ induced by d. It follows that each (generalized) circle A determines its center $z(A)$ and radius uniquely, and that $d(z(A),z(B))\leqq\rho(A,B)$ for all $A,B\in S$.

An Iterated Divisor Function

6064 [1975, 1016]. *Proposed by H. W. Lenstra, Jr., University of Amsterdam, the Netherlands*

For a nonnegative integer m, let $s(m)$ denote the sum of those divisors d of m for which $1 \leqq d < m$. Prove that for every integer $t \geqq 1$ there exists m such that $m < s(m) < s^2(m) < \cdots < s^t(m)$. Here $s^2(m) = s(s(m))$, etc.

Solution by the proposer. More generally, we prove: for every $t \geqq 1$ there exists a positive integer n_t such that $m < s(m) < \cdots < s^t(m)$ for all integers $m \geqq 1$ for which $n_t \| m$. Here $n \| m$ means that m/n is an integer which is relatively prime to n.

The proof is by induction on t. For $t = 1$ we can take $n_t = 12$. Induction step: suppose n_t has the required property; we construct n_{t+1}. Let p be a prime number not dividing n_t, and choose $k \geqq 1$ such that

$$p^k \equiv 1 \pmod{n_t^2 \cdot (p-1)}.$$

(It suffices to take $k = \phi(n_t^2 \cdot (p-1))$ where ϕ is the Euler function.) We claim that $n_{t+1} = n_t \cdot p^{k-1}$ has the required property. To prove this, let m be a positive integer for which $n_{t+1} \| m$. Let $\sigma(m) = s(m) + m$. The function σ is multiplicative, so $p^{k-1} \| m$ implies that $\sigma(m)$ is divisible by $\sigma(p^{k-1}) = (p^k - 1)/(p-1)$, which in turn is divisible by n_t^2. Hence

$$s(m) = \sigma(m) - m \equiv -m \pmod{n_t^2},$$

so $n_t \| m$ implies $n_t \| s(m)$. Therefore

$$s(m) < s(s(m)) < \cdots < s^t(s(m)).$$

Also, $n_t \| m$ implies $m < s(m)$. We conclude $m < s(m) < \cdots < s^{t+1}(m)$ as required.

Reference: H. J. J. te Riele, *A note on the Catalan–Dickson conjecture*. Math. Comp. 27 (1973), 189–192.

The Density of the Sum of Divisors Function

6065 [1975, 1016]. *Proposed by the late C. W. Anderson*

Where $\varphi: N \to N$ is Euler's totient function, it is known that the natural density of $\varphi(N) \subset N$ is zero — in symbols, $d[\varphi(N)] = 0$. Where $\sigma: N \to N$ is the sum of the divisors function, demonstrate that $d[\sigma(N)] = 0$.

Solution by John L. Davison, Laurentian University, Ontario, Canada. In Niven and Zuckerman, *An Introduction to the Theory of Numbers*, is found a proof of the zero density of $\varphi(N)$. A small adaptation of that proof also proves that $d[\sigma(N)] = 0$.

Let $B^k = \{n : n = p_1 \cdots p_r p_{r+1}^{\alpha_1} \cdots p_{r+s}^{\alpha_s} : \alpha_i \geqq 2 \text{ for } 1 \leqq i \leqq s \text{ and } r \leqq k\}$. Then $d(B^k) = 0$. (This is Problem 1, p. 256 of the reference.) If $n \notin B^k$ then n has at least $(k+1)$ prime factors which occur with exponent one and thus $\sigma(n)$ has a factor of 2^k. Let $\varepsilon > 0$, and choose k so that $2^{-k} < \varepsilon/2$. We can write $\sigma(N) = B \cup C$ where $B = \{m \in \sigma(N) : 2^k \mid m\}$, $B \cap C = \emptyset$. Clearly $B(x) \leqq x/2^k < \varepsilon x/2$. Let $C^* = \sigma^{-1}(C)$. Now $n \in C^* \Rightarrow \sigma(n) \in C \Rightarrow \sigma(n)$ has at most k prime factors with exponent one, so $C^* \subset B^k$. Hence $d(C^*) = 0$. But $\sigma(n) > n \vee n$. Hence $C(x) \leqq C^*(x) < \varepsilon x/2$ for x large. Thus $\sigma(N)(x) < \varepsilon x$ for x large and the result follows.

Also solved by J. C. Lagarias, L. E. Mattics, and the proposer.

Note. T. Šalát (Czechoslovakia) informs us that the result of the problem may be found in H. J. Kanold, *Über die Zahlentheoretische Funktionen*, Journal für die Reine und Angewandte Mathematik, 195 (1955), 180–191.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN
with the assistance of the mathematics departments of St. Olaf and Carleton Colleges
COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, CARLETON COLLEGE

We invite readers to submit reviews of significant recent college-level mathematics books. We especially encourage reviews based on classroom use, or comparative reviews of several related books. Reviews should ordinarily not exceed two pages (per book) typed double spaced. Manuscripts of reviews as well as books submitted for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057.

Analysis in Euclidean Space. By Kenneth Hoffman. Prentice-Hall, Englewood Cliffs, New Jersey, 1975. xiv + 432 pp. $16.50 (Telegraphic Review, June–July 1975.)

This text was used in a full-year sequence entitled "Introduction to Analysis" during 1975–76. Students enrolled in this course ranged from sophomores to first year graduate students. The only prerequisite was completion of the traditional calculus sequence, including some familiarity with elementary linear algebra and ordinary differential equations.

The objectives stated at the outset of the course were: to examine the fundamental ideas and basic techniques of analysis, including a review of the concepts of the calculus and their extensions to more abstract settings; to become able to recognize and construct valid mathematical proofs (however intuitive they may be); and to become able to read a sophisticated mathematical text. Both the students and the instructor agreed that the success which was achieved with respect to each of these objectives was attributable, in large measure, to the Hoffman text.

The topics of the first four and one half chapters constitute a somewhat standard selection of topics from advanced calculus, omitting the vector field theory, presented in a very modern fashion. These include a review of the properties of the real number system (emphasis on completeness), R^n, and the complex field; a review of the essentials of linear algebra; a discussion centered around the concepts of convergent sequence and compact set; a study of the usual topics on limits and continuity; and a reexamination of integration and differentiation on the line and in R^n.

Chapter five seems to begin with more of the same (a discussion of real power series), but surprises are in store! Having defined real analytic functions, the author next considers the corresponding definition for functions of a complex variable. The ensuing investigation leads to an illuminating discussion of the central role of Taylor's theorem, and launches an ambitious tour through topics in complex analytic function theory, Fourier series and summability theory, concluding with a discussion of the Dirichlet problem.

The tempo is a bit more relaxed in chapter six, a discussion of normed linear spaces. Chapter seven presents the Lebesgue integral, and the concluding chapter (independent of chapter seven) is a discussion of topics relating to differentiable mappings (inverse and implicit function theorems, change of variable theorem).

This text should be considered seriously for courses whose objectives are similar to those above, for a number of reasons. The text is written with the development of the student, not merely the rigorous presentation of analysis, in mind. The degree of responsibility placed on the reader increases throughout the text, and basic techniques and fundamental concepts are repeatedly emphasized. I was able to observe a distinct maturation process in those students who were seriously involved with the text. Time and again the flavor of what distinguishes analysis from discrete mathematical structures comes pouring through. A typical situation is one in which the author establishes that two functions f and g are equal, not by beginning on one side and computing to the other, but by showing that $\|f-g\|$ is smaller than an arbitrary positive number. Further the exercises are extremely well thought-out and enjoyable, ranging from probing true-false questions, through problems involving straight-forward verifications, to demanding tests of insight and mastery of subtleties. Finally, the range of topics is

broad, as the student is involved in significant applications of linear algebra, elementary properties of matrix-valued functions (including e^{At}), and the elementary theory of complex analytic functions.

Analysis in Euclidean Space is a text for those who wish to communicate to their students the essence, and not just the facts, of analysis.

DENNIS BERKEY, Boston University

FILMS

Adventures in Perception. Produced by Hans Van Gelder, Film Produktie, N.V., The Netherlands. Color and sound, 16mm, 22 minutes. U.S. Release 1973. Available, on free loan, from Royal Netherlands Embassy, 4200 Linnean Avenue, N.W., Washington, DC 20008. Also distributed by BFA Educational Media, 2211 Michigan Avenue, Santa Monica, CA 90404. Purchase $280, rental $22, #1118.

The film *Maurits Escher: Painter of Fantasies* was reviewed in this MONTHLY (V. 83, No. 6, 1976, p. 495). We now have another film about Escher which has special interest for mathematical audiences. I have used both films in my mathematics classes, church groups, and an Escher mini-festival at Albion College. I believe *Adventures in Perception* makes more of an impact on its viewing audience in the way the art works are presented in the film.

Escher has said that if he were ever reincarnated he would want to make films and I believe that this would be the kind of film Escher would make. It has an amusing animation of the *Curl-up*, Escher's response to his own puzzlement as to why Nature never invented the wheel. The film displays many themes: forms, shapes, metamorphosis, convergence and divergence, and cycles, to name a few. There are some fifty works shown using close-ups, cropped shots and full views in an effective cinematic mixture.

Early in the film we see Escher in his studio carving out his woodcut, *Snakes.* The camera takes us to his table to look at many of his polyhedral models and to a view of his tesselated *Sphere with Angels and Devils* and other works. The ribbons of *Cube with Magic Ribbons* are traversed before one sees the entire print and realizes their impossibility. The orthogonality of *Other Worlds* is perceived as we move from panel to panel while the contradictions of *Relativity* are singled out and then offered in totality. We pass through *Day and Night* and experience this change with the moving camera. The impossible conjunction of *Concave and Convex* are seen only after we are presented with their separate reality. Much attention is given to the infinite and perspective as in the joke, *Still Life and Street.* Zoom shots are used to bring out the infinite in *Cubic Space Division, Depth, Castrovala, Circle Limits,* and *Smaller and Smaller.* We are jumping from frog to fish to bird and back in *Verbum* in a very clever treatment of these transformations.

Escher talks about his experiences and discusses his use of Penrose's triangles in the impossible construction of the *Waterfall.* A table is shown cluttered with the now familiar art work of Escher — work which is used on covers and as illustrations for scientists who find his work the only way to convey the joy of their discipline. The final segment is a long, slow scan of *Metamorphose* with only a small piece of the print in view at any given time. This is a marvelous way to display the essence of the print — change. The percussive jazz which accompanies this unfolding is very powerful as is the entire musical score.

I have used this film in my classroom just because it is so rich in images mathematicians and mathematics students love, not necessarily to study, but just to view and enjoy. All classes appreciate a break — what better way to pause and refresh than with this master portrayer of things mathematicians can only think about.

BRIAN J. WINKEL, Albion College

TELEGRAPHIC REVIEWS

Telegraphic reviews are designed to give prompt notice of new books with sufficient information to assist our readers in deciding whether to order an examination copy or to suggest library purchase. Possible uses are indicated as follows:

T = textbook P = professional reading
S = supplementary reading L = undergraduate library purchase
13 to 18 = freshman to second year graduate level usage
1 to 4 = appropriate time in semesters to cover text

Asterisks (*) or question marks (?) denote special positive or negative emphasis, respectively. Pubishers are denoted by standard abbreviations; complete addresses may be found in *Books in Print*.

GENERAL, P, L. *The State of Academic Science: The Universities in the Nation's Research Effort.* Bruce L.R. Smith, Joseph J. Karlesky. Change Mag Pr, 1977, xiv + 250 pp, $5.95 (P). An extensively researched prediction of crisis in academic science, "before the problem emerges as a public issue", due to the present decline in economic and intellectual momentum at the nation's leading research universities. Includes 15 pages specifically devoted to mathematics, based in part on site visits to selected graduate schools; the problems in mathematics are similar to those in other fields, and so is the prospect for serious erosion in the quality of American mathematics research over the next decade. (See summary reports in the *CBMS Newsletter,* 12 (1977) 37, and in the *Chronicle of Higher Education,* May 31, 1977.) LAS

GENERAL, S**(14-16), L***, P. *Fractals: Form, Chance, and Dimension.* Benoit Mandelbrot. Freeman, 1977, xvi + 365 pp, $14.95. A unique enterprise in natural philosophy accompanied by striking computer graphics. Technically, a fractal is a set whose Hausdorff-Besicovitch (or fractal) dimension exceeds its topological dimension; the difference is a measure of irregularity and fragmentation. Familiar fractals include Brownian motions, Cantor sets, and snowflake and Peano curves. But fractals model such diverse phenomena as Swiss cheese, the earth's surface, clusterings of stars, shapes of clouds, windings of rivers and coastlines, turbulence, curdling, liquid crystals, and word frequencies. The book is written in informal essay style; but readers would do well first to read P. Morrison's review of the 1975 French edition (*Scientific American,* 233 (November 1975) 143-144) and M. Gardner's column on it (*Scientific American,* 235 (December 1976) 124-128, 133, 152). PJC

GENERAL, L***. *The VNR Concise Encyclopedia of Mathematics.* Ed: W. Gellert, *et al.* Van N-Rein, 1975, 760 pp, $14.95. A reprinting of *Mathematics at a Glance,* the English translation (published in 1975 by VEB Biblio. Inst., Leipzig) of Kleine Enzyklopädie der Mathematik. See TR, August-September 1976. LAS

GENERAL, P. *Transactions of the Twenty-Second Conference of Army Mathematicians.* US Army Research, Durham, NC, 1976, xix + 613 pp, $16.25 (P). Papers from a May 1976 conference at Watervliet Arsenal, New York. LAS

GENERAL, L**. *The Mathematics Teacher: Cumulative Index, Volume 59-68, 1966-1975.* NCTM, 1976, 90 pp, $5.60 (P). In three parts--by author, by title, and by subject--with complete bibliographic data in each section. LAS

GENERAL, S(13-16), L. *Inside Information, Computers in Fiction.* Ed: Abbe Mowshowitz. A-W, 1977, xxiii + 345 pp, $7.95 (P). Entertaining and provocative "study-anthology" of computers in fiction that focusses on social issues. Pieces are grouped by theme, with each group preceded by a reflective essay by the editor. Extensive bibliography of other fictional works and criticisms, and good indexes. PJC

GENERAL, S(13), L. *Mathematical Puzzles and Perplexities, How to Make the Most of Them.* Claude Birtwistle. Crane, Russak, 1971, 202 pp, $14.50. Collection of puzzles and problems for the intelligent layman. Hints and solutions given (separately). Includes cross-number puzzles for former crossword puzzle addicts. Short bibliography. JG

GENERAL, S*, L*. *Tangrams--330 Puzzles.* Ronald C. Read. Dover, 1965, 152 pp, $1.50 (P). Reprint of a 1965 volume. Hundreds of Tangram patterns, a bit of history, and some variations. Solutions in the back. LAS

GENERAL, S*(9-15), L*. *Tangram, The Ancient Chinese Shapes Game.* Joost Elffers. Trans: R.J. Hollingdale. Penguin, 1976, 169 pp, $5.95 (P). 1,600 shapes, with solutions, supplemented by a brief history, bibliography, and mathematical approach to counting and classifying Tangrams. Comes with a plastic Tangram set. LAS

PRECALCULUS, T(13: 1). *College Algebra, Seventh Edition.* Paul K. Rees, Fred W. Sparks, Charles Sparks Rees. McGraw, 1977, xiv + 487 pp, $13.50. Precalculus book with standard algebra topics, finite math, logarithms and exponentials (but not a functional approach). No trigonometry. Numerous exercises--"essentially all the problems are new." Comprehensive, well written. Many sections have been revised. Level is intermediate for a typical precalculus course. CB

HISTORY. *Carl Friedrich Gauss Werke: Briefwechsel mit F.W. Bessel.* Georg Olms, 1975, xxvi + 597 pp, DM 98. Reprint of 1880 edition of 119 letters exchanged over a 40-year period, almost exclusively devoted to astronomy. A highlight is the appearance in 1829 of the tip of an iceberg: Gauss mentions his 40 years' research in the foundations of geometry and makes the famous remark about fearing "the cries of the Boetians" (p. 490). Bessel's reply indicates that by then non-Euclidean geometry is in the wind (p. 493); Gauss responds with his conviction that "if number is nakedly the product of our mind, space aside from our mind also has a reality, whose laws we cannot completely prescribe *a priori*" (p. 497). PJC

HISTORY, S*(13), P, L*. *Gottfried Wilhelm Leibniz.* Ronald Calinger. Edwin B. Allen Math. Mem., Rensselaer Poly. Inst., 1976, ix + 102 pp, $5 (P). Concise vignette of his life, with brief discussions of his philosophy and his posthumous reputation. Calculus students will enjoy Leibniz's own account of how he came to study mathematics seriously (p. 12). PJC

HISTORY, P, L. *Dirichlet's Principle, A Mathematical Comedy of Errors and its Influence on the Development of Analysis.* A.F. Monna. Oosthoek, Scheltema & Holkema, 1975, vii + 138 pp, $14 (P). This "mixture of history and mathematics" focuses on Dirichlet's method for solving problems arising in potential theory. The author discusses the origins of potential theory, the minimal principles developed by Gauss, Thompson, Dirichlet, and Riemann. The connections between potential theory and function theory, and recent developments in potential theory. A fundamental theme is the consistent confusion of the concepts of maximum (minimum) and upper bound (lower bound). For an extended review see *Historia Mathematica,* 4 (May 1977) 222-225. SG

HISTORY, P, L. *The Mathematical Papers of Isaac Newton, V. VII, 1691-1695.* Ed: D.T. Whiteside. Cambridge U Pr, 1976, xlvii + 706 pp, $95. Papers from Newton's last five Cambridge years, containing his introduction of the dot notation for fluxions, and an explicit treatment of [Taylor] polynomial expansions. More than half the volume is devoted to geometrical studies elaborating on the Greek "topos analuomenos." LAS

HISTORY, S(9-14), L. *The Ages of Mathematics, V. I-IV.* Ed: Charles F. Linn. Doubleday, 1977, $5.95 each. *V. I: The Origins,* Michael Moffatt, 137 pp; *V. II: Mathematics East and West,* Charles F. Linn, 151 pp; *V. III: Western Mathematics Comes of Age,* Cynthia Conwell Cook, 151 pp; *V. IV: The Modern Ages,* Peter D. Cook, 137 pp. An attractively designed set of books offering high-school students (and others) a chronological "non-mathematical survey" of the history of mathematics--that is, emphasis on the history rather than the mathematics. In the large, the books are fine: the style is lively, the tone is right, and the reader is inspired. Numerous illustrations make for easy reading. Vol. II is particularly good in drawing together the isolated strands of mathematics in the Dark Ages. Some inaccurate statements occur (IV, 41: "the validity of Cantor's reasoning is still disputed"; IV, 116: L.F. Richardson (1881-1953) is spoken of as alive and working); but such are not unusual in a popularized treatment. What mars the series, however, is unacceptably poor editing. Two of the illustrations are severly mislabelled (II, 84; and IV, 76, where Babbage looks out over a caption devoted to Dodgson); strange lapses in prose occur in a few places (III, 124; IV, 65) where essential words were cut out by mistake; and various minor errata serve to annoy (II, 73: "sight for "site"; II, 84: "illusions"for "allusions"; IV, 22: "Storm" for "Sturm"; IV, 24: "Lame" for "Lamé"). Since no bibliography or suggestions for further reading are given, one can only guess at the sources used by the authors, who are not professional historians of mathematics but an anthropologist, high-school math teacher, South American historian, and electronics engineer. Despite the flaws, the books will provide students with a valuable social context to humanize the mathematics they encounter. PJC

HISTORY, P. *The Development of Newtonian Optics in England.* Henry John Steffens. Science History Pub, 1977, viii + 190 pp, $12.

HISTORY, S(14-16). *Selected Topics in the History of Mathematics.* Aaron Strauss. U of Maryland, 1975, 107 pp, $3.50 (P). Lecture notes from a one-semester course primarily for math education majors, with a calculus prerequisite: antiquity (18 lectures), renaissance (9), modern (8--more exposition of set theory than history). PJC

FOUNDATIONS, T(17-18: 1). *Einführung in die mathematische Logik.* Hans Hermes. Teubner, Stuttgart, 1972, 206 pp, DM 34 (P). Reprint of third edition (1972) unchanged except for preface. It was the second edition (1969) which was translated into English as *Introduction to Mathematical Logic* (Springer-Verlag, 1973; TR, August-September 1973). Principal difference in third and fourth editions is the greater accent on terminology of model theory in the chapters on Peano arithmetic and the theorems of Robinson, Craig and Beth. PJC

COMBINATORICS, P. *Lecture Notes in Mathematics-560: Combinatorial Mathematics IV.* Ed: Louis R.A. Casse, Walter D. Wallis. Springer-Verlag, 1976, vi + 249 pp, $10.20 (P). Texts of most of the invited lectures and contributed papers from the Fourth Australian Conference on Combinatorial Mathematics in August 1975. JAS

COMBINATORICS, T(13-16: 1), S, L. *Mathematischer Einführungskurs für Informatiker.* Walter Oberschelp, Detlef Wille. Teubner, Stuttgart, 1976, 236 pp, DM 19,80 (P). Topics in discrete structures used in computer science: elementary combinatorics, discrete probability, Boolean algebra, lists, trees, graphs, optimization, codes. Bibliography. Index. RJA

COMBINATORICS, S(11-18), L. *Recursion Sequences.* A.I. Markushevich. Trans: V. Zhitomirsky. MIR (US Dist: Imported Pub, 320 W. Ohio St., Chicago, IL 60610), 1975, 48 pp, $1 (P). "A real though a small mathematical theory": a chapter of the theory of finite differences dealing with complete solution of k-ary linear recurrence relations, developing notions of set of basis sequences, characteristic equation, and general form of solution. Translation needs editing by native speaker of English. PJC

NUMBER THEORY, S(18), P. *Lecture Notes in Mathematics-548: The Selberg Trace Formula for PSL (2,R), V. I.* Dennis A. Hejhal. Springer-Verlag, 1976, vi + 516 pp, $15.20 (P). A comprehensive development of the trace formula for PSL(2,R) in the case of a compact quotient space. The noncompact case will be considered in a subsequent volume. CEC

LINEAR ALGEBRA, T(13-14: 1). *An Introduction to Matrices, Vectors and Linear Programming, Second Edition.* Hugh G. Campbell. P-H, 1977, xiii + 316 pp, $10.95. Changes from the first edition include rewriting and rearrangement of chapters on convex sets, linear programming, vector spaces, and linear transformations and the characteristic value problem. (Shadow prices in linear programming are poorly motivated, and the exposition (p. 218) is marred by misdirected inequalities.) PJC

ALGEBRA, P. *Lecture Notes in Mathematics-549: Brauer Groups.* Ed: D. Zelinsky. Springer-Verlag, 1976, v + 187 pp, $7.40 (P). Proceedings of the conference held at Northwestern University in October 1975. JAS

ALGEBRA, T(14-16: 2), L. *A Survey of Modern Algebra, Fourth Edition.* Garrett Birkhoff, Saunders MacLane. Macmillan, 1977, xi + 500 pp, $13.50. A fine tuning of the 1965 third edition of this classic. The biggest change is in the chapter once called "Algebra of Classes" and now called "Boolean Algebras and Lattices"; however, the basic ideas are still similar in spite of the title change. The fresh looking edition seems to be a timely release for today's modern algebra classes.JAS

ALGEBRA, P. *Les Structures, Nouvelle Pédagogie de la Mathématique.* Remi Ziglon. Hermann (US Rep: SMPF, 111 W. 57th St., NY 10019), 1973, xiv + 219 pp, (P). Topics include sets, groups, rings, fields, vector spaces, and linear maps, all treated at an elementary level. Designed for teachers of college freshmen and sophomores. JG

ALGEBRA, P. *Lecture Notes in Mathematics-551: Algebraic K-Theory.* Ed: Michael R. Stein. Springer-Verlag, 1976, xi + 409 pp, $15.20 (P). Proceedings of the conference held at Northwestern University, January 12-16, 1976. JAS

ALGEBRA, T(17-18: 1), P. *The Rings of Dimension Two.* Wolmer V. Vasconcelos. Lect. Notes in Pure and Appl. Math., V. 22. Dekker, 1976, x + 101 pp, $14.50 (P). Preliminary material on the basic elements of dimension theory is followed by identification and analysis of local rings of global dimension two or less, change of rings, and the study of rings of dimension two which are not necessarily local, particularly rings modulo a finitely generated ideal. The notes conclude with a discussion of coherence of polynomial rings. Note price. JG

ALGEBRA, P. *Lecture Notes in Mathematics-553: Categories of Algebraic Systems.* Mario Petrich. Springer-Verlag, 1976, viii + 217 pp, $10.20 (P). A self-contained exposition of structures common to vector and projective spaces, semigroups, rings and lattices. Should be accessible to anyone with a "general interest in some of these subjects." JAS

ALGEBRA, P. *Lattice Theory.* Ed: A.P. Huhn, E.T. Schmidt. North-Holland, 1976, 462 pp, $48. Submitted papers from the colloquium at Szeged, Hungary in August 1974, together with a program of the colloquium. JAS

FINITE MATHEMATICS, T(13: 1). *Schaum's Outline of Theory and Problems of Discrete Mathematics.* Seymour Lipschutz. McGraw, 1976, 249 pp, $4.95 (P). No probability or linear programming, but a chapter on each of: set theory; relations, functions; vectors and matrices; graph theory; planar graphs, colorations, trees; digraphs and finite state machines; combinatorics; algebraic systems and formal languages; posets and lattices; propositional calculus; and Boolean algebra. PJC

CALCULUS, T(13: 3). *Calculus, Second Edition.* Lynn Loomis. A-W, 1977, xiii + 888 pp, $18.95. Major changes from the first edition (TR, January 1975) include completely rewritten chapters on the definite integral, inverse functions, and vectors; there is a new chapter on Green's Theorem. The problem sets have been reorganized, with non-routine problems collected with other miscellaneous exercises at the end of each chapter. The presentation is more cohesive--the material on curve sketching is now in one chapter, and the definite integral notation has been moved out of the early chapter on antidifferentiation. JG

CALCULUS, T(13: 3). *College Calculus with Analytic Geometry, Third Edition.* Murray H. Protter, Charles D. Morrey, Jr. A-W, 1977, ix + 849 pp, $18.95. Essentially the same as second edition (TR, August 1970). Intuitive definition of limits is given early, but ε-δ approach is used in remainder of text. New chapter on solid analytic geometry, with unified treatment of vectors in two and three dimensions in a separate chapter. Additional more challenging exercises. JG

CALCULUS, T???(13). *Calculus for the Life Science, An Introduction.* Murray A. Katz. Dekker, 1976, xii + 258 pp, $13.75. Written by a physician for students of medicine and biology in order to "enable [them] to develop a facility for handling the calculus necessary for research." The author calls it a "cookbook", yet feels compelled to derive many results. Invariably the "derivations" are flawed or false! One can only hope *his* doctor doesn't learn from this book! TAV

REAL ANALYSIS, T(14-15: 1, 2). *A First Course in Real Analysis.* M.H. Protter, C.B. Morrey. Springer-Verlag, 1977, xii + 507 pp, $18.80. A comprehensive development of the theory behind single and multi-variable calculus, set forth in clear, spare theorem-proof-remark prose. There is more than enough material for a full year post-(or, the authors suggest, honors-level) calculus course. One might wish, however, for more examples and more motivating intuition. LAS

REAL ANALYSIS, T(15-16: 1), L. *Mathematical Analysis.* Gabriel Klambauer. Pure and Appl. Math., V. 31. Dekker, 1975, viii + 500 pp, $24.50. From the construction of the reals using Dedekind cuts, through sequences, series, continuity, differentiation, integration, uniform convergence to metric spaces, the reader encounters a tightly written text. Contains abundant exercises, many with hints or solutions. A possible flaw: very few figures and little motivational material is included to give the reader an intuitive feeling for the result. TAV

DIFFERENTIAL EQUATIONS, P. *Lecture Notes in Mathematics-561: Function Theoretic Methods for Partial Differential Equations.* Ed: V.E. Meister, N. Weck, W.L. Wendland. Springer-Verlag, 1976, xviii + 520 pp, $17.60 (P). The proceedings, less papers "to be published elsewhere", of an April 1976 conference at the Technische Hochschule, Darmstadt, Germany. JAS

DIFFERENTIAL EQUATIONS, P. *Asymptotic Expansions for Ordinary Differential Equations.* Wolfgang Wasow. Krieger, 1976, ix + 374 pp, $20. A survey which concentrates on mathematics methods in the subject rather than on specific equations. The author discusses those methods which lead to full infinite expansions. Among the topics covered are turning point problems, nonlinear equations, and singular perturbations. SG

DIFFERENTIAL EQUATIONS, P. *Singularly Perturbed Differential Operators of Second Order.* P.P.N. de Groen. Math. Centre Tracts, No. 68. Math Centrum, 1976, ix + 159 pp, Dfl. 20 (P). A study of spectral properties and asymptotics of certain singularly perturbed two-point boundary value problems on (-1,1). SG

DIFFERENTIAL EQUATIONS, P. *Lecture Notes in Mathematics-564: Ordinary and Partial Differential Equations.* Ed: W.N. Everitt, B.D. Sleeman. Springer-Verlag, 1976, xviii + 551 pp, $18.50 (P). Proceedings of the fourth conference held at Dundee, Scotland on March 30 to April 2, 1976. JAS

NUMERICAL ANALYSIS, P. *Lecture Notes in Mathematics-556: Approximation Theory.* Ed: R. Schaback, K. Scherer. Springer-Verlag, 1976, vii + 466 pp, $15.20 (P). Proceedings of the international colloquium held in June 1976 at the Institut für Angewandte Mathematik der Universität Bonn, Germany.JAS

NUMERICAL ANALYSIS, P. *Lecture Notes in Physics-59: Proceedings of the Fifth International Conference on Numerical Methods in Fluid Dynamics.* Ed: A.I. van de Vooren, P.J. Zandbergen. Springer-Verlag, 1976, vii + 459 pp, $15.20 (P). Complete proceedings of the conference held June 28 to July 3, 1976 at Twente University of Technology in Enschede, The Netherlands. JAS

NUMERICAL ANALYSIS, T(15: 1), S. *Computational Mathematics, Worked Examples and Problems with Elements of Theory.* N.V. Kopchenova, I.A. Maron. MIR, 1975, 395 pp. A handbook of numerical methods intended for engineering students. Includes numerous methods for function evaluation, interpolation, linear and nonlinear systems, differentiation and integration, initial and boundary value problems for ordinary differential equations, integral equations and partial differential equations. The discussions of the methods are generally brief. A good selection of examples and problems. RWN

NUMERICAL ANALYSIS, T(15-16: 1), S. *Einführung in die numerische Mathematik.* Eduard Stiefel. Teubner, Stuttgart, 1976, 292 pp, DM 24,80 (P). The fifth edition of an introduction to numerical analysis written for engineers, physicists and mathematicians. Differs from the fourth largely in having chapters on spline interpolation and the method of finite elements. JD-B

FUNCTIONAL ANALYSIS, P. *Lecture Notes in Mathematics-519: Espaces de Fonctions Continues.* Jean Schmets. Springer-Verlag, 1976, xii + 150 pp, $7.40 (P). An exposition and generalization of the author's recent work. The author studies spaces of continuous functions and characterizes spaces associated to these function spaces. SG

FUNCTIONAL ANALYSIS, P. *On the Feynman Integral.* Klaus Brock. Aarhus U, 1976, 49 pp, (P). A description of both Wiener and Feynman integrals in terms of families of linear functionals and a limiting process due to K. Ito. JAS

FUNCTIONAL ANALYSIS, P. *Lecture Notes in Mathematics-563: Séminaire de Théorie du Potentiel Paris, No. 2.* Ed: M. Brelot, G. Choquet, J. Deny. Springer-Verlag, 1976, vi + 292 pp, $10.20 (P). Some of the papers from the 1975-1976 seminar. JAS

OPTIMIZATION, T, P. *Programmation Linéaire.* Claude Guérard. Pr U Montreal, 1976, 416 pp, $13.25 (P). An introduction to linear programming. Discusses linear equations and inequalities, the simplex method, duality, etc. A special feature is the inclusion of several interesting case studies. SG

OPTIMIZATION, S*(14-16), L*. *The Shortest Lines, Variational Problems.* L.A. Lyusternik. Trans: Yuri Ermolyev. MIR (U.S. Distr: Imported Pub, 320 W. Ohio St., Chicago, IL 60610), 1976, 104 pp, $1 (P). An elementary, semi-popular introduction to maximization (and minimization) problems depending on the shape of a curve. Treats geodesics, potential energy in stretched threads, the isoperimetric problem, and Fermat's principle. LAS

OPTIMIZATION, T(16-18: 2), P, L*. *Differential Games, A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization.* Rufus Isaacs. Krieger, 1975, xxiv + 398 pp, $19. Reprinting of 1965 original Wiley edition, preceded by a substantive new introduction and supplemented by notes and new references. The original monograph was based in large part on the author's original work for the Rand Corp. LAS

OPTIMIZATION, T(15-17: 1, 2), L. *Linear and Combinatorial Programming.* Katta G. Murty. Wiley, 1976, xxiii + 567 pp, $18.95. Thorough study of linear and integer programming for undergraduates and graduates in a variety of fields; linear algebra prerequisite. Besides the essential discussion of the simplex method and its variants, the author attends to computational efficiency and numerical stability; other topics include parametric linear programs, sensitivity analysis, network algorithms, integer programming methods, branch and bound, and the complementary pivot algorithm (for linear programs, quadratic programs, and bimatrix games). PJC

ALGEBRAIC GEOMETRY, P. *Curves and Their Jacobians.* David Mumford. U of Michigan Pr, 1976, 104 pp, $5 (P). The simplicity of the title belies the setting, which is high-level algebraic geometry. Mumford brilliantly brings out the interrelation of analytic, algebraic and geometric techniques.PJC

GEOMETRY, T(15). *Groupes de Lie, Représentations Linéaires et Applications.* Guy Pichon. Hermann (US Distr: SMPF, 111 W. 57th St., NY 10019), 1973, 142 pp, (P). After a quick survey of topological groups, the author presents the basics of compact groups. A solid background in algebra, topology and geometry is presumed, although the relevant definitions and theorems are stated when needed. SG

GEOMETRY, S(10-12). *Mathematics Through Paper Folding.* Alton T. Olson. NCTM, 1975, iv + 60 pp, $1.75 (P). Classical geometry (constructions, theorems), polygon geometry, some algebra, and standard recreations all exhibited by clever folding of paper. A revision of Donovan Johnson's *Paper Folding for the Mathematics Class.* LAS

PROBABILITY, P. *Stochastic Approximation and Recursive Estimation.* M.B. Nevel'son, R.Z. Has'minskii. Trans. Math. Mono., V. 47. AMS, 1973, iv + 244 pp, $30.40. Devoted to sequential methods for solving problems such as location of zeros or maxima for functions when the measured value of the function contains a random error. Primarily concerned with stochastic approximation using Markov processes and martingales. The translation is very smooth. Extensive bibliography with mostly Russian sources. TAV

STATISTICS, T(13-15: 1), S. *Introduction to Statistical Analysis, A Semiprogrammed Approach.* Celeste McCollough. McGraw, 1974, xvi + 380 pp, $8.95 (P). Paperback edition of hardcover original (TR, October 1975), $1 cheaper. LAS

COMPUTER SCIENCE, P. *Computer Science and Multiple-Valued Logic, Theory and Applications.* Ed: David C. Rine. North-Holland, 1977, xiv + 548 pp, $50.95. Demand for more sophisticated and reliable computers with greater speed in less volume has prompted investigation of multiple-valued logic as an alternative logic because it works with greater information density. The areas in which logic functions are used (program flow control, operations on logical variables, and operations on bit strings) do not inherently favor binary structures, but binary hardware has offered the greatest reliability. This volume investigates algebraic theory, logic design and switching theory, threshold logic design, physical components and implementations. Recommended surveys on pp. 81-101, 475-484. (Pp. 501-516 were missing from the review copy.) PJC

COMPUTER SCIENCE, P. *Sixth Annual ACM Symposium on Theory of Computing.* ACM, 1974, iv + 347 pp, $11 (P).

COMPUTER SCIENCE, T(13: 1). *Programming Standard COBOL.* Winchung A. Chai, Henry W. Chai. Acad Pr, 1976, xviii + 342 pp, $9.95 (P). A gradual and systematic introduction to programming in ANSI Cobol. Numerous sample programs. A chapter on structural programming. Explains hardware dependent features for several systems although assumes 360/370's. Exercises. RWN

COMPUTER SCIENCE, T(13-14: 1), S. *Structured Programming in PL/I and PL/C.* Bernhard Fischer, Herman Fischer. Dekker, 1976, xi + 402 pp, $14.75. Teaches PL/I emphasizing the structured programming features of the language. Includes many sample programs with complete explanations of statements used. Many descriptive diagrams and tables. Comprehensive treatment does justice to the many sophisticated capabilities of PL/I. Many problems. Chapter quizzes. Answers to quizzes and problems. Eight appendices. Bibliography. Index. RJA

COMPUTER SCIENCE, T(16-17: 1), S, P. *Algorithmentheorie.* Jacques Loeckx. Springer-Verlag, 1976, xiv + 223 pp, $11.50 (P). An introduction to the theory of algorithms, intended for students of information theory rather than of logic. Deals, among other topics, with Turing machines, recursive functions, and nondeterministic algorithms. Exercises, many with solutions or hints. JD-B

COMPUTER SCIENCE, P. *Computer Science and Scientific Computing.* Ed: James M. Ortega. Acad Pr, 1976, x + 306 pp, $15.50. Scientific computing spans many disciplines. Included are thirteen papers (complete with references) written by experts from various areas involved in scientific computing. Presented at the Third ICASE Conference on Scientific Computing. RJA

COMPUTER SCIENCE, T(13-16: 1), S, L. *Computer System Architecture.* M. Morris Mano. P-H, 1976, xiv + 478 pp, $17.95. A computer system includes both hardware and software. This text emphasizes the hardware, yet machine level software is discussed where appropriate. First part deals with data and the hard/software features that manipulate it. Second part provides an in-depth discussion of the separate functional units of a computer. Well organized. Problems. Chapter references. Index. RJA

COMPUTER SCIENCE, S(15-18). *Certificate in Data Processing Examination.* James W. Morrison. ARCO, 1976, 640 pp, $12 (P). Contains a study guide with review questions and answers. Intended for those studying for the Annual Certificate in Data Processing Examination. Includes three practice exams with answers. RJA

COMPUTER SCIENCE, T(17-18: 1), S, P. *Algebraic and Automata-Theoretic Properties of Formal Languages.* Seymour Ginsburg. North-Holland, 1975, xii + 313 pp, $27.75. Begins with general mathematical and language notions and discussion of the Chomsky hierarchy of languages and automata. The major portion of the text is devoted to the study of the structures "trio", "semi-AFL", and "AFL", along with the basic transformation device, "a-transducer." Exercises. References. Indices. RJA

COMPUTER SCIENCE, T(17-18: 1, 2), P. *Lecture Notes in Computer Science-36: Theory of Program Structures: Schemes, Semantics, Verification.* Sheila A. Greibach. Springer-Verlag, 1975, xv + 364 pp, $13.20 (P). Very readable treatment for computer science graduate students of the theory of abstract flowcharts and program verification, including decision problems. Requires elementary knowledge of formal languages, automata, and logic. Some exercises gathered at end. PJC

COMPUTER SCIENCE, P. *Lecture Notes in Computer Science-39: Data Base Systems.* Ed: H. Hasselmeier, W.G. Spruth. Springer-Verlag, 1976, vi + 386 pp, $13.20 (P). Proceedings of the Fifth Informatik Symposium held in Bad Homburg, Germany in September 1975. JAS

COMPUTER SCIENCE, P. *Sequencing by Enumerative Methods.* J.K. Lenstra. Math. Centre Tracts, No. 69. Math Centrum, 1977, x + 198 pp, Dfl. 24 (P). After a discussion of various NP-complete problems, focussing on complexity of machine scheduling problems, the author turns to enumerative methods and a recursive approach to their implementation. This revised doctoral thesis then concludes with branch and bound algorithms and practical applications. PJC

COMPUTER SCIENCE, S(16-17), P. *Arithmetik in Rechenanlagen, Logik und Entwurf.* Otto Spaniol. Teubner, Stuttgart, 1976, 208 pp, DM 24,80 (P). Descriptions and comparisons of fast and cost-efficient algorithms for performing arithmetic operations on digital computers. Deals with pipelining, the CORDIC technique and time-complexity, as well as more basic matters. JD-B

COMPUTER SCIENCE, T**(15-17: 1), S, L. *Computer Data-Base Organization.* James Martin. P-H, 1975, xviii + 558 pp, $26.50. Written by one of the experts in the field. First half on the logical organization of data bases; second part on the physical organization. Filled with helpful diagrams and graphs. Detailed but well organized. References. Appendices. Indices. Chapter questions. RJA

COMPUTER SCIENCE, T(14-16: 1), S, L*. *Adaptive Information Processing, An Introductory Survey.* Jeffrey R. Sampson. Springer-Verlag, 1976, x + 214 pp, $14.80. Wide-ranging but elementary survey of diverse topics (information and automata, biological information processing, artificial intelligence), keyed to annotated bibliographies of sources and accompanied by a small number of exercises. PJC

COMPUTER SCIENCE, P. *Lecture Notes in Computer Science-33: Automata Theory and Formal Languages; 2nd GI Conference.* Ed: H. Brakhage. Springer-Verlag, 1975, viii + 292 pp, $11.50 (P). Proceedings of the conference held in May 1975 at Kaiserslauten, Germany. JAS

APPLICATIONS (BUSINESS), T(15-17: 1), S, L*. *Introduction to Decision Theory.* J. Morgan Jones. Irwin, 1977, xiv + 369 pp, $15.50. Introductory text in decision theory for the "mathematically naive student" (prerequisite: algebra; probability is developed as needed). Excellent extended analysis of expected value criterion; main focus is on sampling strategies (especially binomial sampling problems). The reader is befriended with an introductory rationale for studying the subject, an informal style throughout, and a concluding summary of what has and has not been learned. (An instructor's manual, not sent with review copy of text, describes relevant computer programs in APL and Basic and how to obtain and use them.) PJC

APPLICATIONS (CARTOGRAPHY), P. *Conformal Projections Based on Elliptic Functions.* L.P. Lee. Cartographica Monograph, No. 16. U of Toronto Pr, 1976, 128 pp, $4 (P). The elliptic functions of Dixon and Jacobi turn out to be the natural tools for effecting transverse Mercator projections of the sphere. Included are details of conformal projections of the sphere upon the five regular polyhedra (cf. Buckminster Fuller's *Dymaxion World Map*). Other notable features include historical perspective, orientation toward calculation (much use of series and approximations), and finely drafted figures.PJC

APPLICATIONS (CARTOGRAPHY), P. *The Seven Aspects of a General Map Projection.* Thomas Wray. Cartographica Monograph, No. 11. U of Toronto Pr, 1974, 72 pp, $4 (P). Formerly, three aspects were distinguished for conical projections; Wray extends this to 7 for a general map projection. The underlying mathematical idea is the symmetry group of the projection; the author works with fundamental regions, with which aspects are associated much as crystal classes are with three-dimensional fundamental regions. Errata sheet. PJC

APPLICATIONS (COMMUNICATION), P. *Digital Pattern Recognition.* Ed: K.S. Fu. Springer-Verlag, 1976, xi + 206 pp, $29.80. Summary of major recent developments in pattern recognition. Divided into six chapters, each of which is written by experts in the subject. Expensive. Subject index. RJA

APPLICATIONS (CONTROL THEORY), P. *Control Theory and Topics in Functional Analysis, V. I-III.* Abdus Salam. Int. Atomic Energy Agency, Vienna, Austria (U.S. Distr: UNIPUB, Box 433, Murray Hill Station, NY 10016), 1976. *V. I*, 463 pp; *V. II*, 319 pp; *V. III*, 419 pp, $73 (P). The lectures presented at the seminar course in Trieste from September 11 to November 29, 1974. JAS

APPLICATIONS (ECONOMICS), P. *Lecture Notes in Economics and Mathematical Systems-135: A Disequilibrium-Equilibrium Model with Money and Bonds, A Keynesian-Walrasian Synthesis.* Hanjiro Haga. Springer-Verlag, 1976, vi + 119 pp, $7.40 (P).

APPLICATIONS (ECONOMICS), P**, L*. *Values of Non-Atomic Games.* R.J. Aumann, L.S. Shapley. Princeton U Pr, 1974, xi + 333 pp, $16. An extension of the classical "Shapley value" of a finite game to infinite games in which no individual player has any influence; includes applications to traditional notions such as the core of a game and economic equilibrium. An axiomatic, highly mathematical treatment, closely related to non-atomic measure theory. LAS

APPLICATIONS (GENETICS), P, L. *Foundations of Mathematical Genetics.* A.W.F. Edwards. Cambridge U Pr, 1977, viii + 119 pp, $13.50. Basic deterministic models for random-mating diploid populations with constant viabilities; populations are assumed large enough for stochastic variation to be neglected. Presumes familiarity with calculus and linear algebra. Most of the theory has been developed in the last 10-20 years; the author provides historical notes and an extensive bibliography. PJC

Reviewers Whose Initials Appear Above

Richard J. Allen, St. Olaf; Cecelia Bleecker, Carleton; Paul J. Campbell, St. Olaf; Clifton E. Corzatt, St. Olaf; John Dyer-Bennet, Carleton; Jennifer Galovich, St. Olaf; Steven Galovich, Carleton; R.W. Nau, Carleton; J. Arthur Seebach, St. Olaf; Lynn Arthur Steen, St. Olaf; T.A. Vessey, St. Olaf.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, NW, Washington, D. C. 20036. Items must be submitted at least five months before publication can take place.

PERSONAL ITEMS

Albion College: Associate Professor Ronald Fryxell has been promoted to Professor; Assistant Professor John Wenzel has been promoted to Associate Professor.

Professor Marion Beiter, O.S.F., Daemen College, has retired with the title of Professor Emeritus after 30 years of service to the college.

Lt. Col. R. L. Eisenman is a specialist in National Defense in the Library of Congress while continuing as vice-president of Ferry Landing Woods and as lecturer at the University of Maryland.

Dr. Ruth S. Lefkowitz retired from the Mathematics Department at John Jay College of Criminal Justice, CUNY, on September 30, 1976, with the title of Professor Emeritus. She served as the first chairman of the Department of Mathematics which was established at the college in 1972.

Professor C. Douglas Olds, San Jose State University, has retired with the title of Professor Emeritus.

Instructor R. V. Olson, Worcester Polytechnic Institute, has been promoted to Assistant Professor.

Dr. G. M. Wing, Los Alamos Scientific Laboratory, has been appointed Professor and Chairman of the Department of Mathematics at Southern Methodist University.

Professor Paul Brock, University of Vermont, died on February 2, 1977, at the age of 53. He was a member of the Association for thirty-five years.

Professor Paul A. Clement, Washington State University, died on November 16, 1976, at the age of 60. He was a member of the Association for twenty-eight years.

Assistant Professor Sue King Dunkle, Clemson University, died on September 27, 1976, at the age of 63. She was a member of the Association for fourteen years.

Mr. Rodney Espey, Portland Community College, died on August 5, 1976, at the age of 36. He was a member of the Association for nine years.

Dr. Fred G. Fender, Rutgers University, died on May 24, 1976, at the age of 67. He was a member of the Association for twenty-one years.

Professor Walter H. Leser, Franklin and Marshall College, died on January 27, 1977, at the age of 51. He was a member of the Association for twenty-three years.

Professor M. P. O'Donnell, University of Queensland, Australia, died on November 10, 1976, at the age of 49. He was a member of the Association for twenty-three years.

Professor Tullio J. Pignani, East Carolina University, died on September 30, 1976, at the age of 56. He was a member of the Association for twenty-six years.

Professor Emeritus Susan M. Rambo, Smith College, died on January 7, 1977, at the age of 93. She was a member of the Association for fifty-nine years.

Dr. Charles R. Sherer, Fort Worth, Texas, died on February 6, 1977, at the age of 82. He was a member of the Association for fifty-two years.

Professor James Singer, Brooklyn College, CUNY, died on June 17, 1976, at the age of 70. He was a member of the Association for thirty-eight years.

Mrs. M. T. Wolbier, Board of Education, Buffalo, New York, died on October 15, 1976, at the age of 58. She was a member of the Association for eighteen years.

Professor Emeritus F. Lynwood Wren, California State University, Northridge, died on October 20, 1976, at the age of 82. He was a member of the Association for forty-nine years. The University has established the F. Lynwood Wren Scholarship Fund in his memory. Contributions may be sent to: The University Foundation, California State University, Northridge, California 91330.

1977–78 CHAUTAUQUA-TYPE SHORT COURSES FOR COLLEGE TEACHERS

Come October the 1977–78 series of NSF Chautauqua-Type Short Courses for College Teachers will offer the opportunity for some 3,500 college teachers to participate. Approximately fifty short courses will be offered in the 1977–78 academic year at field centers throughout the U.S. The program is designed to enable college teachers to

keep abreast of advances in a variety of fields of science and to help them incorporate these advances in their teaching. The program is conducted by the AAAS with support from the National Science Foundation. Further information is available from the AAAS. A preliminary announcement poster, available in May, lists the course titles and course directors. A program announcement brochure, available in July, gives full details — course descriptions, course directors, field center locations, schedules, eligibility and application forms. Please write to: American Association for the Advancement of Science, Office of Science Education, Box C22, 1776 Massachusetts Avenue, N.W., Washington, D.C. 20036.

THE 1977–78 SABBATICAL EXCHANGE INFORMATION SERVICE

For the past two years, the MAA Sabbatical Exchange Information Service (SEIS) has provided information to assist faculty members in universities and colleges (both two- and four-year) in arranging what might be called "no-cost sabbaticals." Reaction from participants has been enthusiastic, but SEIS remains under-used. Since the value of such an information exchange to its users increases with the number of listings, we are making an effort to broaden the visibility and hence the scope of SEIS. We ask readers not only to consider using SEIS, but also to remind colleagues of its existence. It costs nothing to be listed in SEIS, and the results can be a year of expanded horizons at little or no cost to faculty member or institution. Here is how SEIS operates.

Many MAA members are in institutions not offering faculty members a program of sabbatical leaves. Even in institutions with sabbatical leave programs individual faculty members often find themselves ineligible for such a leave at a time when the desire for one is strongest. We all recognize the rejuvenating effect of an occasional change of scene, even if it does not involve release from teaching duties. The Association therefore suggests that an occasional exchange between two faculty members of similar interests, training, and experience at different institutions could be of great benefit to the individuals and also to their institutions. The individuals and institutions all stand to gain from the refreshment of exchanged ideas and insights.

It is often possible for two such faculty members to trade identities, so to speak, for a year. Such an exchange might involve trading teaching responsibilities, living quarters, and some departmental responsibilities. The extent of the exchange would depend on the individual circumstances. It is suggested, however, that salaries should not be exchanged or even discussed. Each faculty member would remain on the payroll of his permanent institution and receive all of his normal fringe benefits. Financially, his institution would not recognize the exchange at all.

The MAA proposes to become involved only to the extent of assisting in bringing together like-minded mathematics faculty members who are interested in an exchange. The information exchange will be accomplished by the annual publication by the Association in December of a list containing the names, addresses, and other pertinent information about members of the Association interested in arranging a "Sabbatical Exchange" with a colleague in another institution. This list will be sent free of charge to all those on the list and to any other MAA member who requests it.

Members interested in being listed in December 1977 should write to "SEIS, The Mathematical Association of America, 1225 Connecticut Avenue, N.W., Washington, D.C. 20036," enclosing the following information about themselves:

1. Name
2. Institution
3. Department
4. Address
5. Rank
6. Major field of interest
7. Highest earned degree
8. Names of from one to five courses recently taught
9. Normal teaching load
10. Section of country preferred for visit: Northeast, Southeast, Northcentral, Southcentral, Northwest, Southwest.
11. Period for which exchange is desired, e.g., all of the academic year 1978–79, or the first two quarters of 1978–79, or the second semester of 1978–79, etc.

Communications must reach the Washington office by November 19, 1977, for inclusion in the December 1977 list.

THE COOPERATIVE COLLEGE REGISTER

The Cooperative College Register has been re-established as a communications link and matching service for positions and position-seekers for higher education. Write for details: Cooperative College Register, 621 Duke Street, P. O. Box 298–A, Alexandria, Virginia 22314.

AUSTRALIAN MATHEMATICAL SOCIETY

An Applied Mathematics Conference sponsored by the Division of Applied Mathematics, Australian Mathematical Society, will be held on February 5 to 8, 1978, at the Broadbeach Hotel, Gold Coast, Queensland. A session on the teaching of Applied Mathematics is to be included. For details write to Dr. R. D. Braddock, Department of Mathematics, University of Queensland, St. Lucia, 4067, Queensland, Australia.

THE AUSTRALASIAN MATHEMATICAL CONVENTION

The Australasian Mathematical Convention will run from May 15 to 19, 1978. It will bring together the Australian and New Zealand Mathematical Societies and Associations for the first time. Enquiries should be sent to: 1978 Convention Secretary, Department of Mathematics, University of Canterbury, Christchurch, 1, New Zealand.

THE GREATER METROPOLITAN NEW YORK MATH FAIR

THE GREATER METROPOLITAN NEW YORK MATH FAIR will be held March 5, 1978; papers from high-school students in or near New York city will be due January 13, 1978. Further details and application forms may be obtained from: Professor John Chiaramonte, MATH FAIR Committee, Dept. of Mathematics & Computer Science, St. John's University, Jamaica, New York 11439.

MATH-POEMS

Primary Press (E. Robson, Box 105, Parker Ford, PA 19457) is soliciting "Math-poems" for a projected book. Constraints: "1. Poems should include either numbers, curves, equations, tables, diagrams or geometric figures. 2. An interesting subjective verbalization that may be humorous or emotional in any way so long as it is imaginative and/or interpretive. No money in this for anyone. Only fun, fame, fantasy and, maybe, beauty?"

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

THE 1977 WILLIAM LOWELL PUTNAM MATHEMATICAL COMPETITION

The 38th annual William Lowell Putnam Mathematical Competition will be held at participating institutions on Saturday December 3, 1977. This competition is supported by the William Lowell Putnam Prize Fund For The Promotion of Scholarship and is administered by The Mathematical Association of America. All colleges and universities in Canada and the United States may register eligible undergraduates. Registration forms will be mailed to schools that participated in either the 36th or the 37th competition by September 15, 1977. Other institutions that wish to enter undergraduates should request registration material from Dr. A. P. Hillman, Director; The William Lowell Putnam Mathematical Competition; 709 Solano Dr., S. E.; Albuquerque, New Mexico, 87108. Completed registrations must be received by the director no later than October 28, 1977.

Further details are given in the *Announcement Brochure* that is mailed with the registration blanks. Reports of previous competitions, including examination questions and outlines of solutions, are in past issues of this MONTHLY; the most recent of these reports were in the issues of Nov. 1976, Nov. 1975, Dec. 1974, and Nov. 1973.

NEW SECTIONAL GOVERNORS OF THE ASSOCIATION

The following have been elected Governors of the Association representing the Sections indicated:

FLORIDA	C. W. McArthur, Florida State University
ILLINOIS	Jon M. Laible, Eastern Illinois University
INTERMOUNTAIN	William J. Coles, University of Utah
IOWA	James L. Cornette, Iowa State University
LOUISIANA–MISSISSIPPI	Thomas A. Atchison, Mississippi State University
MARYLAND–DC–VIRGINIA	Theodore J. Benac, U. S. Naval Academy
MICHIGAN	Yousef Alavi, Western Michigan University
NORTH CENTRAL	Warren S. Loud, University of Minnesota
PHILADELPHIA	Jerry P. King, Lehigh University
SOUTHERN CALIFORNIA	Alicia A. Huffman, California State Polytechnic University
TEXAS	James N. Younglove, University of Houston

The highest percentage of voters was 43%, occurring in the Iowa Section. The Intermountain Section was the runner-up with 36%.

A. B. WILLCOX, *Executive Director*

SUGGESTION BOX

Members of the MAA are encouraged to send in suggestions, questions, etc., about the operations of the Association. Communications will be referred to the appropriate officer of the Association for answering; from time to time, those of general interest may also be answered in one or both of the official journals. Communications should be addressed to: Suggestion Box, Mathematical Association of America, 1225 Connecticut Avenue, N. W., Washington, D. C. 20036.

NOVEMBER MEETING OF THE INDIANA SECTION

The fall meeting of the Indiana Section of the MAA was held at Manchester College, North Manchester, on Saturday, November 6, 1976, with approximately 60 persons in attendance. The Chairman of the Section, M. C. Gemignani, IUPU–at Indianapolis, presided.

The program consisted of the following:

1. *Remedial mathematics: an administrator's viewpoint*, by M. C. Gemignani, IUPUI.
2. *A glimpse of algebraic number theory, parts I and II*, by B. Pollak, Notre Dame.
3. *Panel discussion on the new State Bulletin concerning Teacher Education.* Panelists: E. Alton, IUPUI; D. Deal, Ball State; P. Nugent, Franklin College.

After the panel discussion the Section took the following actions:

Passed. *Resolution I*: Those electing the science area will be eligible to teach mathematics only if 24 hours are completed in mathematics equivalent to the mathematics minor.

Passed. *Resolution II*: A provision shall be made for a computer science endorsement that can be added to a license with a mathematics or science teaching major. This endorsement would permit a teacher to teach computer science, computer language, and computer literacy courses.

Endorsed. Recommendations of The Indiana Mathematics Educators for Teacher Certification Programs regarding mathematics requirements.

D. E. WILSON, *Secretary*

NOVEMBER MEETING OF THE NORTHEASTERN SECTION

The twenty-second annual meeting of the Northeastern Section of the MAA was held at Rhode Island College, Providence, Rhode Island, on November 27, 1976; there were 91 people in attendance. The section chairman, Grattan Murphy, presided.

The morning meeting was devoted to a panel discussion: *High School Preparation for College Mathematics Courses*, Moderator: R. D. Klein, Northeastern University; panelists: Arthur Bardige, Arlington, Massachusetts

Public School System; G. S. Cunningham, University of Maine at Orono; Romualdas Skvarcius, Institute for Curriculum Development, Boston University.

At the afternoon business meeting the following officers were elected for the coming year: Chairman, E. C. Schlesinger, Connecticut College; Vice-Chairman, D. B. Small, Colby College; Secretary-Treasurer, G. W. Best, Phillips Academy. Chairman Murphy announced the appointment of the following regional MAA contest chairmen: Atlantic provinces, William Crawford; Maine, Alfred Harper, Jr.; Massachusetts and Rhode Island, Murray Abramson; Vermont, James Burgmeier. A brief description of the successful June meeting at the University of New Hampshire on the topic of Industrial Mathematics was presented by Chairman Murphy. The business meeting concluded with Donald Small's report on the 1975 High School Lecture Program.

The following talks completed the program:

Prime generating functions and congruences, by H. L. Alder, President Elect of the MAA, University of California, Davis.

A theory of island zoogeography, by W. H. Bossert, Gordon MacKay Professor of Applied Mathematics, Harvard University.

GEORGE BEST, *Secretary-Treasurer*

FEBRUARY MEETING OF THE NORTHERN CALIFORNIA SECTION

The 1977 meeting of the Northern California Section of the MAA was held at San Francisco State University on February 26, 1977. There were 138 persons in attendance.

The following invited addresses were presented:

1. *Lattice points and convex sets*, by Don Chakerian, University of California, Davis.
2. *Rational reciprocity laws*, by Emma Lehmer, Berkeley, California.
3. *Non-cancellation phenomena in arithmetic and topology*, by Peter Hilton, Battelle Institute, Seattle, Washington.
4. *Tone perception and decomposition of periodic functions*, by David Gale, University of California, Berkeley.

A luncheon panel discussion on "Dispelling Math Anxiety and Getting Ready for Calculus Fast" was attended by 70 persons. Jane Day, College of Notre Dame, served as moderator and panelists included: Lenore Blum, Mills College; Karel De Leeuw and Phil Faillace, Stanford University.

The business meeting was presided over by Section Chairperson David Barnette, University of California, Davis. The following officers were elected for 1977–1978: Chairperson, Jane Day, College of Notre Dame, Belmont; Vice-Chairperson, Herbert Holden, Stanford Research Institute, Menlo Park; Program Chairperson, David Barnette, University of California, Davis; Secretary-Treasurer, Newman Fisher, San Francisco State University.

NEWMAN FISHER, *Secretary-Treasurer*

MARCH MEETING OF THE FLORIDA SECTION

The Tenth Annual Spring Meeting of the Florida Section of the MAA was held on March 4 and 5, 1977, at the University of South Florida in Tampa, Florida.

Six invited addresses were presented as follows: "Prime Generating Functions and Congruences," Professor H. L. Alder, University of California, Davis, California, and President of MAA; "The Meaningless Calculus Course," Professor Morris Kline, New York University; "How to Tell that a Simple Overhand Knot is Really Knotted," Professor Edwin Moise, Queen's College; "On Teaching Teachers (or, Trig. Made Simple)," Professor M. P. Hale, Jr., University of Florida; "Lattice Graphs and Applications," Professor F. O. Hadlock, Florida Atlantic University; "The 3rd International Congress on Mathematical Education," Professor Don Hill, Florida A & M University.

In conjunction with the meeting there was a State Articulation Conference. The following talks were presented to the Conference: "Implementation of Hand Calculators in Undergraduate Classes," A. D. Snider, University of South Florida; "The Future Role of Mathematics Articulation in Florida," panel discussion involving everyone, Bill Rice, presiding; "Believe It or Not—It is a Record," Ignacio Bello, Hillsborough Community College; "The Journals Pertaining to Mathematics Education," E. J. Bolduc, University of Florida.

A Saturday morning session was sponsored by Pi Mu Epsilon, the Mathematics Honorary Fraternity. The following talks were presented: "Introduction to Infinitesimal Calculus," Joni Hersch, University of South Florida; "Analytic Functions and Elementary Particles," Mario Pita, University of South Florida; "Pi Mu Epsilon in the State of Florida," J. S. Frame, Michigan State University.

The following papers were presented to the section:

1. *On the number of vertices in a grid graph and related problems*, by Roy Levow, Florida Atlantic University.
2. *Equational completeness*, by John Leeson, University of North Florida.
3. *A different ordering in certain function spaces*, by Nancy Fordyce, Florida State University.

4. *Plane trigonometry for an arbitrary quadratic space*, by B. H. Edwards, University of Florida.
5. *Opposite and reciprocal in learning algebra*, by Alan Wayne, Pasco-Hernando Community College.
6. *An algorithm for linear programming*, by J. S. Frame, Michigan State University.
7. *Prediction in polygenic systems*, by Allan Anderson, City University of New York.
8. *On a conjecture of Lorentz*, by Michael Lachance, University of South Florida.
9. *A Markov chain model for studying the effect of remedial math work*, by J. Higgins and J. Schwenke, University of South Florida.
10. *Computer programs*, by W. H. Green, John F. Kennedy Space Center.
11. *The optimal shape of a tent*, by R. C. Jones, Jr., Florida Technological University.
12. *The use of computer generated modules in Undergraduate Mathematics Instructions*, by E. P. Miles, Jr., Florida State University.

The luncheon-business meeting was held Saturday, March 7, 1977. Chairman A. W. Goodman presided at the meeting. Committee reports were presented and Professor George Lofquist of Eckerd College was elected chairman-elect; Professors David Sherry of the University of West Florida and Thomas Ribley of Valencia Community College, were elected as Vice-Chairmen.

F. L. CLEAVER, *Secretary*

MARCH MEETING OF THE SOUTHERN CALIFORNIA SECTION

The Southern California Section of the MAA held its fifty-seventh annual Spring meeting at Loyola Marymount University, Los Angeles, California, on March 12, 1977. One hundred thirty-four attended the meeting. The Chairman of the Section, Professor Paul Yale of Pomona College, presided.

At the business meeting, the results of the election of section officers for 1977–78 were announced: Chairman, John Todd, California Institute of Technology; 1st Vice-Chairman, James Murphy, California State College, San Bernardino; 2nd Vice-Chairman, H. A. Dekleine, California State Polytechnic University, San Luis Obispo; Program Chairman, D. G. Babbitt, University of California, Los Angeles. Paul Yale, Pomona College, becomes Past Chairman, and Edmund Deaton, San Diego State University, continues as Secretary-Treasurer.

The luncheon speaker was Professor John Todd; his talk was entitled "Oberwolfach 1945." Local arrangements were made by Professor Wade Peterson, Loyola Marymount University.

The following program was presented:

A new numerical analysis of the ancient astronomical observations, by Paul Muller, Jet Propulsion Laboratory.

Panel Discussion: *Pre-College Mathematics in College.* Alice Huffman, California State Polytechnic University, Pomona (Moderator); David Cohen, UCLA; Michael Cullen, Loyola Marymount University; Wanda Marosz, San Diego State University; Jack Wadhams, Golden West College.

Calculator demonstrations for a first calculus course, by George McCarty, University of California, Irvine.

A new paradigm for the mathematics classroom, by Alvin White, Harvey Mudd College.

A guided tour through the applications of generalized matrix inverses, by Victor Lovass–Nagy, CSU Fullerton and Clarkson College.

The four color theorem, by Kenneth Holladay, California Institute of Technology.

Einstein's theory and the finiteness of the universe, by Greg Galloway, Loyola Marymount University.

Mathematical models in drug theory, by Greg Pierce, CSU Fullerton.

The Association for Women in Mathematics presented a panel discussion on recruitment of women into mathematics and mathematics anxiety in women. The panelists were Jacqueline Dewar, Loyola Marymount University; Ruth Afflack, California State University, Long Beach; Janet Williams, University of California, Irvine; Susan Cohen, Westchester High School.

E. I. DEATON,, *Secretary-Treasurer*

APRIL MEETING OF THE SOUTHEASTERN SECTION

The fifty-sixth annual meeting of the Southeastern Section was held on April 1–2, 1977, at the University of Alabama in Huntsville, Alabama. A total of 269 persons attended the meeting, including 210 members of the Association. The local arrangements were handled by P. G. Casazza.

Three invited addresses were given: Lida K. Barrett (Section Lecturer) of the University of Tennessee at Knoxville on "Stereology and Mathematics"; H. L. Alder (President of the Association) of the University of California at Davis on "Recent Developments in the Theory of Partition Identities"; and C. C. Lindner of Auburn University on "How to Imbed a Steiner Triple System." There was a Symposium on Mathematics in the Two-Year Colleges, with presentations by J. E. Bright, P. Capell, N. Fentress, J. W. Guest, and W. R. Wilson.

There were seven sessions for contributed papers. The presiders were E. V. Haynsworth (Chairman of the Section), P. G. Casazza, R. A. Dobyns, and I. C. Gentry for the general sessions; and C. A. Brown, A. L. Hudson,

B. G. Klein, H. V. Park, D. J. Pokrass, S. C. Ross, and J. R. Wall for the special sessions. The University hosted a dinner party for attendees on Friday night.

Officers elected for 1977–78 were: Chairman, I. C. Gentry, Wake Forest University; Chairman-elect, John Kenelly, Clemson University; Vice-Chairman, J. E. Cicero, Clayton Junior College; and Section Lecturer, J. H. Carruth of the University of Tennessee at Knoxville.

At the business meeting, the winner of the $25 prize from the Section for the best performance on the Putnam Examination in the Section was announced to be J. B. Zipperer of Armstrong State College. The Section voted to hold its 1979 meeting at the University of Tennessee at Chattanooga.

The following papers were presented:

1. *Computer program to write presentations of all groups of order p^{m+1} with elementary abelian subgroup of order p^{m+1}*, by Thorna Humphries, Bennett College.
2. *Illustration of a covering theorem for the semigroup T_x*, by Bette Donahay and John Nichols, Maryville College.
3. *Characterization and combinatorics of particular inverse subsemigroups of T_x*, by Kenneth Dickens, Maryville College.
4. *A BASIC program to factor $\lambda^p - 1$ in $z_q[\lambda]$*, by Veronica Watson, Bennett College.
5. *The shooting method for existence of solutions to two-point boundary value problems*, by Sarah E. Brown, Wake Forest University.
6. *Real and complex trigonometric identities via uniqueness theorems*, by B. G. Klein, Davidson College.
7. *Indeterminate forms of exponential type*, by J. V. Baxley and E. K. Hayashi, Wake Forest University.
8. *A qualitative study of the two dimensional linear homogeneous differential system*, by T. A. Newton, Washington State University and Clemson University.
9. *Distortion properties of α-spirallike functions and some related classes of analytic functions*, by C. A. Brown, University of Tennessee at Chattanooga.
10. *On some asymptotic formulae of S. Wigert*, by D. Suryanarayana, Memphis State University.
11. *New methods for teaching signed number operations in developmental mathematics*, by Aleeta P. Christian, Roane State Community College.
12. *On small group instruction in a remedial mathematics program*, by P. T. Chang and J. P. Downes, Georgia State University.
13. *A computer-based laboratory for general statistics*, by R. S. Cunningham, Birmingham-Southern College.
14. *Hand-held calculators in the mathematics classroom*, by J. R. Shipman, Alabama A. and M. University.
15. *The power method for determining the maximum eigenvalue for a matrix*, by Claudia Mitchell and E. D. Dixon, Tennessee Technological University.
16. *Minicomputers and the theory of equations*, by Mark Hyde, Lenoir-Rhyne College.
17. *A computer program to write presentations of all the groups of order $q^m p$ with normal elementary abelian Sylow q-subgroup*, by Barbara Johnson, Bennett College.
18. *Groups of order $q^m p$ with elementary abelian Sylow q-subgroups*, by James Alonso, Bennett College.
19. *A class of functions containing* exp, sine, cosine, sinh *and* cosh *as special cases*, by R. E. Worth and F. A. Massey, Georgia State University.
20. *k-linearly independent subsets of finite vector spaces*, by J. K. Doyle and K. I. Mandelberg, Emory University.
21. *Solving $dx/dt = Ax(t) + f(t)$ and $dx/dt = (pI + A)x(t) + f(t)$ when $A^2 = kA$ for some scalar k*, by R. E. Worth and F. A. Massey, Georgia State University.
22. *Cayley tables for dihedral subgroups*, by Donald Cook, Albany Junior College.
23. *A placement system for freshman mathematics*, by L. M. Perry and L. Kitchens, Appalachian State University.
24. *Description of a large scale mastery format for basic algebra*, by H. T. Matthews, University of Tennessee at Knoxville.
25. *Advantages and disadvantages of a large scale mastery format for basic algebra*, by W. E. Haver, University of Tennessee at Knoxville.
26. *Mastery learning applied to mathematics methods at Shelby State Community College*, by Gerald Hagerman, Shelby State Community College.
27. *On coordinating the lower level multiple sections mathematics courses*, by Satish Chandra, Tuskegee Institute.
28. *A computer assignment for power series*, by E. W. Womble and M. M. Galloway, Presbyterian College.
29. *Hand-held computers and the numerical methods course*, by G. M. Eargle, Appalachian State University.
30. *The coconut problem and the hand calculator*, by D. R. Stone, Georgia Southern College.
31. *A guide for job hunting in the area of computer science*, by T. F. Higginbotham, Auburn University.
32. *The centro-involution of a matrix*, by J. J. Johnson, University of Tennessee at Martin.

J. D. NEFF, *Secretary-Treasurer*

CALENDAR OF FUTURE MEETINGS

Sixty-first Annual Meeting, Atlanta, Georgia, January 6–8, 1978.

Fifty-eighth Summer Meeting, Brown University, Providence, August 8–10, 1978.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN, University of Pittsburgh, Pennsylvania, April 21–22, 1978, or April 28–29, 1978.

FLORIDA, St. Petersburg Junior College, Clearwater, March 3–4, 1978.

ILLINOIS, Western Illinois University, Macomb, May 5–6, 1978.

INDIANA, Indiana Central College, Indianapolis, November 5, 1977.

INTERMOUNTAIN

IOWA, University of Northern Iowa, Iowa Falls, April 22, 1978.

KANSAS, Wichita State University, Wichita, late March–early April 1978.

KENTUCKY, early April. Deadline for papers 6 wks. bef. mtg.

LOUISIANA-MISSISSIPPI, Buena Vista Hotel-Motel, Biloxi, Mississippi, February 17–18, 1978.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA, American University, Washington, D.C., November 19, 1977.

METROPOLITAN NEW YORK, late April or early May 1978. Deadline for papers 2 wks. bef. mtg.

MICHIGAN, Michigan State University, East Lansing, Spring 1978.

MISSOURI, Central Missouri State University, Warrensburg, April 7–8, 1978.

NEBRASKA, University of Nebraska at Omaha, April 14–15, 1978.

NEW JERSEY, Caldwell College, Caldwell, November 5, 1977.

NORTH CENTRAL, University of Minnesota, Morris, October 14–15, 1977.

NORTHEASTERN, Saturday after Thanksgiving, and third week in June in odd-numbered years.

NORTHERN CALIFORNIA, College of Notre Dame, Belmont, February 1978.

OHIO, Wright State University, Dayton, October 28–29, 1977.

OKLAHOMA-ARKANSAS, Henderson State University, Arkadelphia, Arkansas, March 31–April 1, 1978.

PACIFIC NORTHWEST, second Saturday in June. Deadline for papers 6 wks. bef. mtg.

PHILADELPHIA, Moravian College, Bethlehem, November 19, 1977.

ROCKY MOUNTAIN, South Dakota School of Mines and Technology, Rapid City, April 28–29, 1978.

SEAWAY, SUNY College at Plattsburgh, New York, October 28–29, 1977.

SOUTHEASTERN, Clemson University, Clemson, South Carolina, March 31–April 1, 1978.

SOUTHERN CALIFORNIA, California State Polytechnic University, San Luis Obispo, November 11–12, 1977.

SOUTHWESTERN, New Mexico Institute of Mining and Technology, Socorro, Spring 1978.

TEXAS, Stephen F. Austin State University, Nacogdoches, March 31–April 1, 1978.

WISCONSIN, University of Wisconsin—Whitewater, late April 1978.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Washington, February 12–17, 1978.

AMERICAN MATHEMATICAL ASSOCIATION OF TWO YEAR COLLEGES, Peachtree Plaza Hotel, Atlanta, Georgia, October 13–14, 1977.

AMERICAN MATHEMATICAL SOCIETY, Atlanta, Georgia, January 4–7, 1978.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION

ASSOCIATION FOR COMPUTING MACHINERY, Olympic Hotel, Seattle, Washington, October 17–19, 1977.

ASSOCIATION FOR SYMBOLIC LOGIC

ASSOCIATION FOR WOMEN IN MATHEMATICS, Atlanta, Georgia, January 4–8, 1978.

CANADIAN SOCIETY FOR HISTORY AND PHILOSOPHY OF MATHEMATICS/SOCIÉTÉ CANADIENNE D'HISTOIRE ET DE PHILOSOPHIE DES MATHÉMATIQUES

FIBONACCI ASSOCIATION

INSTITUTE OF MATHEMATICAL STATISTICS

MU ALPHA THETA

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, San Diego, California, April 12–15, 1978.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Peachtree Plaza Hotel, Atlanta, November 7–9, 1977.

PI MU EPSILON

SCHOOL SCIENCE AND MATHEMATICS ASSOCIATION, William Penn Hotel, Pittsburgh, Pennsylvania, November 10–12, 1977.

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, Hilton Inn, Albuquerque, New Mexico, October 31–November 2, 1977.

Eminent Mathematicians and Mathematical Expositors speak to

STUDENTS and TEACHERS in. . .

The NEW MATHEMATICAL LIBRARY

An internationally acclaimed paperback series providing •

- *stimulating excursions for students beyond traditional school mathematics*
- *supplementary reading for school and college classrooms*
- *valuable background reading for teachers*
- *challenging problems for solvers of all ages from high school competitions in the US and abroad.*

The New Mathematical Library is published by the MATHEMATICAL ASSOCIATION OF AMERICA. The volumes are paperbound.

LIST PRICE: $4.50 each volume. Price to MAA MEMBERS AND HIGH SCHOOL STUDENTS (Prepaid only): $3.50 each volume.

NUMBERS: RATIONAL AND IRRATIONAL by Ivan Niven, NML-01

WHAT IS CALCULUS ABOUT? by W. W. Sawyer, NML-02

AN INTRODUCTION TO INEQUALITIES, by E. F. Beckenbach, and R. Bellman, NML-03

GEOMETRIC INEQUALITIES, by N. D. Kazarinoff, NML-04

THE CONTEST PROBLEM BOOK. Problems from the Annual High School Mathematics Contests sponsored by the MAA, NCTM, Mu Alpha Theta, The Society of Actuaries, and the Casualty Actuarial Society. Covers the period 1950-1960. Compiled and with solutions by C. T. Salkind. NML-05

THE LORE OF LARGE NUMBERS, by P. J. Davis, NML-06

USES OF INFINITY, by Leo Zippin, NML-07

GEOMETRIC TRANSFORMATIONS, by I. M. Yaglom, translated by Allen Shields, NML-08

CONTINUED FRACTIONS, by C. D. Olds, NML-09

GRAPHS AND THEIR USES, by Oystein Ore, NML-10

HUNGARIAN PROBLEM BOOKS I and II, based on the Eotvos Competitions 1894-1905 and 1906-1928. Translated by E. Rapaport, NML-11 and NML-12

EPISODES FROM THE EARLY HISTORY OF MATHEMATICS, by A. Aaboe, NML-13

GROUPS AND THEIR GRAPHS, by I. Grossman and W. Magnus, NML-14

THE MATHEMATICS OF CHOICE, by Ivan Niven, NML-15

FROM PYTHAGORAS TO EINSTEIN, by K. O. Friedrichs, NML-16

THE MAA PROBLEM BOOK II. A continuation of NML-05 containing problems and solutions from the Annual High-School Mathematics Contests for the period 1961-1965. NML-17

FIRST CONCEPTS OF TOPOLOGY, by W. G. Chinn and N.E. Steenrod, NML-18

GEOMETRY REVISITED, by H.S.M. Coxeter, and S. L. Greitzer, NML-19

INVITATION TO NUMBER THEORY, by Oystein Ore, NML-20

GEOMETRIC TRANSFORMATIONS II, by I. M. Yaglom, translated by Allen Shields, NML-21

ELEMENTARY CRYPTANALYSIS—A Mathematical Approach, by Abraham Sinkov, NML-22

INGENUITY IN MATHEMATICS, by Ross Honsberger, NML-23

GEOMETRIC TRANSFORMATIONS III, by I. M. Yaglom, translated by Abe Shenitzer, NML-24

THE MAA PROBLEM BOOK III. A continuation of NML-05 and NML-17, containing problems and solutions from the Annual High School Mathematics Contests for the period 1966-1972. NML-25

MATHEMATICAL METHODS IN SCIENCE, by George Pólya, NML-26

Send orders to: **The Mathematical Association of America**
1225 Connecticut Ave., NW, Washington, D.C. 20036

THE MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.
Washington, DC 20036

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

AN OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 84 NUMBER 8

CODEN: AMMYAE

CONTENTS

(*Continued on inside cover*)

NOTICE TO AUTHORS

Research papers per se are unsuitable; see statement of policy (Vol. 81, p. 1).

Please follow the format in current issues of the Monthly. Manuscripts must be legibly typewritten or reproduced from typewritten copy, double spaced with wide margins and on one side of the paper. Three copies should be submitted to the appropriate editor and one kept by the author as protection against loss. The author's full address *must* appear at the end of the manuscript.

Backlog: Main Articles 24 months, Math. Notes 17 months, Research Problems 16 months, Classroom Notes 14 months, Math. Education 17 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to R. P. Boas, Department of Mathematics, Northwestern University, Evanston, IL 60201; NOTES, etc.: to the corresponding Associate Editor; REPRINT PERMISSION: to Leonard Gillman, Mathematical Association of America, University of Texas, Austin, Texas 78712 (see also the copyright notice below); ADVERTISING CORRESPONDENCE: to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo, Buffalo, N.Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. Willcox, Mathematical Association of America, 1225 Connecticut Ave., N. W., Washington, D.C. 20036; BACK ISSUES: Contact P. and H. Bliss Co., Middletown, CT 06457.

Annual dues for members of the Association, including a subscription to the American Mathematical Monthly, are \$18.00 for each of the first two years of membership and \$22.00 thereafter. Student Membership is available with annual dues of \$12.00. For nonmembers the subscription price is \$25.00.

Published by the Association at Washington, D.C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D.C., and additional mailing offices.

THE LOGIC OF EQUALITY

LEON HENKIN[1]

Try asking a high school student if he can give you any equations which hold identically in the system $\mathfrak{N} = (N, +, 0)$ of natural numbers under addition. If you formulate the question using terminology adjusted to his textbook, there's a good chance, these days, that he'll answer "Sure," and quote — without thinking — the commutative and associative laws and the "zero law," $x + 0 = x$ for all $x \in N$. He doesn't have to think because his teacher has probably required him to memorize these CAZ laws just as, in the old days, we all had to memorize addition and multiplication tables.

Now ask the student if he can give you any *other* laws, besides those three, which hold identically in the system $\mathfrak{N}$ — and you may provoke some thought. If the thinking is successful, the student will come up with some equation like $x + y = y + (0 + x)$, and he will be able to prove that it holds identically in the system $\mathfrak{N}$ by deriving it from the CAZ identities, using rules of equational logic.

Go one more step, now, and ask if the student can give you an equation that holds identically in $\mathfrak{N}$ and *which cannot be derived from the CAZ laws.* This time you have probably passed beyond his depth. Indeed, this question is likely to provoke uneasiness among a good many mathematicians. True, most mathematicians will quickly guess that there is *no* equational identity for $\mathfrak{N}$ which cannot be derived from the CAZ laws, on the grounds that any such law would have been discovered long ago and that they would surely have heard of it. But they will be a little uneasy at the thought of trying to prove this conjecture, when they realize that such a proof must be based on a mathematically precise definition of what it means for one identity to be derivable from others. Although mathematicians can generally distinguish a correct derivation from an incorrect one, and can construct derivations (of both kinds), it is only the specialist in logic who feels comfortable at the prospect of giving — and using — a definition of the set of all valid derivations.

The rules of proof (or derivation) that one needs in a given context depend upon the form of the sentences with which one is dealing. In the case where we restrict ourselves to sentences having the form of equational identities, the rules can be formulated very simply, and indeed are pretty widely known. Because of this simplicity, the logic of equational identities lends itself particularly well to an exposition of such fundamental metamathematical notions as *soundness* and *completeness*, which involve the distinction and relation between truth and provability. It is the purpose of this paper to describe these matters, and then to return to the question of identities in the system $\mathfrak{N}$ and related systems.

Before we begin, however, let us take one more look at the high school textbook. Thank goodness, it still treats the solution of linear and quadratic equations! When a student solves such an equation by traditional methods, he is not only finding the solutions but also, in essence, constructing a proof that these are the only solutions. This proof makes use of equational identities that hold in the underlying number system, but it also uses logical rules of proof appropriate for other equations, such as the given one, which do *not* hold identically. The rules of proof for these two types of equations are so similar that it is easy to overlook the need for a distinction; in Section 7 below we shall distinguish and treat both of these types of equations.

1. Equational grammars. Any formal deductive system consists of two parts. First, a classified list of symbols is provided and rules are given for constructing formulas from them; second, rules of inference are described, enabling us to pass from given formulas to others. The first of these two components may be thought of as a kind of grammar for an idealized language.

[1] This work was supported in part by the National Science Foundation, Grant MPS74–23878. The content of the paper formed the subject of an invited address at the annual meeting of the MAA held in San Francisco in January, 1974. The author is grateful to the referee, as well as to Don Pigozzi, who read an earlier version of the paper and made useful suggestions.

The grammars with which one deals in mathematical logic themselves possess a well defined mathematical structure. This, to a certain degree, mirrors the mathematical structure of the domains which are intended as the subject matter for interpreted languages based upon the grammars. The similarity of structure between a grammar and its subject domains facilitates the use of language for communication, and also plays a key role in completeness proofs for deductive systems based upon the grammar. Accordingly, we shall give emphasis to these structural ideas in the description which follows of equational grammars.

Let us consider in detail a grammar G designed to refer to an arbitrary structure having the same type as $\mathfrak{R}$, that is, to any system $\mathfrak{A} = (A, +_A, 0_A)$ where A is a non-empty set, $+_A$ is a 2-place operation on A, and 0_A is an element of A.[2] The *symbols* of the grammar G are to include an operation symbol $+$, an "individual constant" $\mathbf{0}$, left and right parentheses, (and), an "identity sign" $\equiv$, and a further set V of symbols $v_0, v_1, \ldots$, called "variables."

Having listed the symbols of G, we now have to consider strings of these symbols and to select from among all such strings those which are to be called (equational) *identities*. To save time here we may simply identify strings with finite sequences of symbols, and we shall suppose known the operation $^\frown$ of concatenation which, when applied to strings σ and τ, produces a new string $\sigma^\frown\tau$ whose symbols consist of those of σ "followed by" those of τ.[3]

Let S be the set of all strings of symbols of G. In order to define the subset of S consisting of identities, we need first to define an auxiliary subset T of S whose elements will be called *terms*; these terms will serve as the "sides" of the identities. Strings consisting of a variable v_i of V, or of the individual constant $\mathbf{0}$, alone, will be terms; and further terms will be built up from these by means of the following binary operation $+_S$ on strings: For any $\sigma, \tau \in S$ we set

$$\sigma +_S \tau = (^\frown\sigma^\frown + ^\frown\tau^\frown).^4$$

Thus T is defined as the intersection of all those subsets $U \subseteq S$ such that (i) $V \cup \{\mathbf{0}\} \subseteq U$, and (ii) U is closed under the operation $+_S$. Clearly T is itself one of these sets U, and indeed it is the least such set.

Let $+_T$ be the restriction of $+_S$ to T, so that $+_T$ is a 2-place operation on T. Since $\mathbf{0} \in T$, we see that

(1.1) $\mathfrak{T} = (T, +_T, \mathbf{0})$ is a *structure*.

The structure $\mathfrak{T}$ has the following properties:

(1.2) *The sets* V, $\{\mathbf{0}\}$, *and* **range** $(+_T)$ *are pairwise disjoint subsets of* T.

(1.3) *If* $\sigma, \tau, \sigma', \tau' \in T$ *and* $\sigma +_T \tau = \sigma' +_T \tau'$, *then* $\sigma = \sigma'$ *and* $\tau = \tau'$.

(1.4) *If* $U \subseteq T$ *and* (i) $V \cup \{0\} \subseteq U$ *and* (ii) U *is closed under* $+_T$, *then* $U = T$.[5]

The structure $\mathfrak{T}$ is called *the algebra of terms* of the grammar G. It turns out that the conditions (1.1)–(1.4) characterize $\mathfrak{T}$ to within isomorphism; hence in the sequel there is never any need to revert

[2] Henceforth we shall use the word *structure* alone to refer to a structure having the same type as $\mathfrak{R}$, unless an indication of some other type is explicitly given.

[3] An alternative procedure is to consider the set S of strings as an undefined set (having the symbols of G among its elements), the operation $^\frown$ is taken as an undefined operation on S, and then suitable axioms are provided as a basis for the theory of strings.

[4] Here we have tacitly identified a symbol, such as + or), with the string having that symbol as its only component. We continue this practice in the sequel.

[5] Only the proof of (1.3), called "the unique readability condition," presents any difficulty. (Note that the condition fails in S, i.e., we can have $\sigma, \tau, \sigma', \tau' \in S$ with $\sigma +_S \tau = \sigma' +_S \tau'$ but $\sigma \neq \sigma'$ and $\tau \neq \tau'$; for instance, put $\sigma = \mathbf{0}$, $\tau = (^\frown)^\frown + ^\frown(^\frown\mathbf{0}^\frown)$, $\sigma' = \mathbf{0}^\frown + ^\frown(^\frown)$, and $\tau' = (^\frown\mathbf{0}^\frown)$.) To prove (1.3) we show that every term has the same number of components (as it has components), while if $\sigma \in$ *range* $(+_T)$ then every proper initial string of σ must contain more components (than the number of its components).

to the definitions of T and $+_T$, as all arguments involving terms of G can be based on properties (1.1)–(1.4).

To complete the description of our grammar G we specify that an *identity* is a string of the form $\sigma^\frown \equiv {}^\frown\tau$, where σ, τ are any terms. Henceforth we shorten our notation and write simply $\sigma \equiv \tau$ for identities, and we also call an identity a *formula* of G.

2. Equational deductions. *Deductions* (also called *derivations*) are finite sequences of formulas leading from a given set Γ of formulas, the *premises* of the deduction, to another formula ϕ, the *conclusion* of the deduction. Formulas in a deduction which are not premises must be related to earlier formulas of the deduction by specified relations called *rules of inference.* The grammar of formulas together with the rules defining deductions make up a *formal deductive system.*

We now present a formal deductive system based upon the grammar G of Section 1.

Rules of inference for the grammar G.

(R1) To infer an identity $\sigma \equiv \tau$ from the identity $\tau \equiv \sigma$, where σ and τ are any terms.[6] (Rule of *symmetry.*)

(R2) To infer an identity $\sigma \equiv \tau$ from identities $\sigma \equiv \rho$ and $\rho \equiv \tau$. (*Rule of transitivity.*)

(R3) To infer an identity $(\sigma +_T \sigma') \equiv (\tau +_T \tau')$ from identities $\sigma \equiv \tau$ and $\sigma' \equiv \tau'$. (*Rule of replacement.*)

(R4) To infer $(\mathrm{Sub}_f \sigma) \equiv (\mathrm{Sub}_f \tau)$ from the identity $\sigma \equiv \tau$, where $\sigma, \tau \in T$, f is any map of V into T, and $\mathrm{Sub}_f \sigma$, $\mathrm{Sub}_f \tau$ are the terms obtained respectively from σ, τ by substituting, for each occurrence of any variable v, the term $f(v)$. (*Rule of substitution.*)[7]

In addition to the rules of inference above, we define a set of formulas of G which will play the role of logical axioms in our deductive system. These formulas are called *tautologies.*

(R0) A formula $\sigma \equiv \sigma$, where σ is any term, is called a *tautology.*[8]

Now for any set Γ of formulas of G we define a *deduction from* Γ to be a finite sequence of formulas $(\phi_1, \ldots, \phi_n)$ such that each ϕ_i is either an element of Γ, or is a tautology, or else can be inferred from one or two earlier formulas of the sequence by one of the rules (R1)–(R4). For any set Γ of formulas, and any formula ϕ, we say that Γ *yields* ϕ, or that ϕ *is derivable* (*deducible*) *from* Γ, iff there is a deduction from Γ whose last formula is ϕ. *Notation*: We write $\Gamma \vdash \phi$ to indicate that Γ yields ϕ.

Within the deductive system $(G, \vdash)$ defined above, we can carry out various derivations of the kind discussed in the opening paragraphs of this paper. For example, we can formulate the CAZ laws by the following identities ϕ_C, ϕ_A, ϕ_Z:

---

[6] We have formulated this rule using the traditional terminological mode "To infer ... from —." From the purely mathematical viewpoint we are simply defining a binary relation R_1 which holds for a pair (ϕ, ψ) of identities iff $\phi = \sigma \equiv \tau$ and $\psi = \tau \equiv \sigma$ for some $\sigma, \tau \in T$. In formulating the remaining rules we shall take for granted that σ, τ, ρ are arbitrary terms without explicit mention. Rules (R2) and (R3) define ternary relations R_2 and R_3 on the set of all identities, while Rule (R4), like (R1), defines a binary relation.

[7] The operation of substituting given terms for variables throughout another given term is intuitively familiar. The operation can be defined precisely by using the definition of terms as selected strings (i.e., sequences) of symbols. An indication of how substitution may be defined when terms are characterized by properties (1.1)–(1.4) of Section 1, is given in Section 3 below.

[8] Although the specification of axioms, as in (R0), is not generally considered to define a rule of inference, it is possible to regard such a set of axioms as constituting a zero-premise rule. Just as the one-premise rules (R1) and (R4) define binary relations on the set of all identities, and the two-premise rules (R2) and (R3) define ternary relations, so the "zero-premise rule" (R0) defines a "unary relation", or subset, of the set of all identities.

$$
\begin{aligned}
&\phi_C = (v_0 +_T v_1) \equiv (v_1 +_T v_0),\\
(2.1)\qquad &\phi_A = ((v_0 +_T v_1) +_T v_2) \equiv (v_0 +_T (v_1 +_T v_2)),\\
&\phi_Z = (v_0 +_T \mathbf{0}) \equiv v_0,
\end{aligned}
$$

we can set $\Gamma_{CAZ} = \{\sigma_C, \sigma_A, \sigma_Z\}$, and we can then establish that

$$\Gamma_{CAZ} \vdash (v_0 +_T v_1) \equiv (v_1 +_T (\mathbf{0} +_T v_0))$$

by producing a suitable deduction from Γ_{CAZ} whose last formula is $(v_0 +_T v_1) \equiv (v_1 +_T (\mathbf{0} +_T v_0))$. (There is one such deduction containing 8 formulas.)

For an example of a general result about the relation $\vdash$, we cite the fact that whenever Γ is *empty* and $\Gamma \vdash \sigma \equiv \tau$, then we must have $\sigma = \tau$. This follows easily from the definition of $\vdash$ by checking that whenever a formula $\pi \equiv \rho$ is inferred from one or from two tautologies by one of the rules (R1)–(R4), then $\pi \equiv \rho$ must itself be a tautology.

If a deductive system is *sound*, the rules of inference should never lead us from premises which are true to a conclusion which is false. Our deductive system $(G, \vdash)$ is, indeed, sound, but in order to prove this we must first have a mathematically precise definition of the notions of truth and falsity for formulas of the grammar G. This matter will be discussed in the following section. The concepts introduced there will lead in Section 4 to a semantical notion of *consequence*, and we can then take up the question whether every formula which is a consequence of a given set Γ of formulas, is derivable from Γ by our rules of inference. This is the question of *completeness* for the deductive system $(G, \vdash)$.

We conclude the present section by remarking that there are alternative definitions of the deductive system $(G, \vdash)$. In particular, (R0) may be weakened by taking only the single axiom $v_0 \equiv v_0$. And (R1)–(R3) may be replaced by a single rule of "replacing equals by equals": To infer $\sigma' \equiv \tau'$ from $\sigma \equiv \tau$ and $\pi \equiv \rho$, where σ' (resp., τ') is either σ (resp., τ) itself, or else is obtained from it by replacing one part from the form π, by a part of the form ρ.

3. Interpretation of G. Consider the formula ϕ_C of the previous section which expresses the commutative law. Is it true or false? Obviously, the answer depends on which structure we are talking about. If the grammar G is being used as the basis for a language referring to the system of natural numbers $\mathfrak{N} = (N, +, 0)$, then ϕ_C becomes a true sentence — we express this by the notation $W^{\mathfrak{N}}(\phi_C) = \mathsf{T}$. On the other hand if exp is the 2-place exponential operation on N, $\exp(x, y) = x^y$, and if we interpret G as referring to the structure $\mathfrak{N}^* = (N, \exp, 0)$, then ϕ_C becomes a false sentence — $W^{\mathfrak{N}^*}(\phi_0) = \mathsf{F}$ — since, e.g., we have $\exp(2,3) \neq \exp(3,2)$. In general, to every identity ψ of G, and every structure $\mathfrak{A} = (A, +_A, 0_A)$, we will associate a truth-value $W^{\mathfrak{A}}(\psi)$, either T or F. One of our aims in this section is to give a precise definition of W.

Of course in a rough way we can say that if ψ is the identity $\sigma \equiv \tau$, then $W^{\mathfrak{A}}(\psi) = \mathsf{T}$ iff the terms σ and τ have the same value in $\mathfrak{A}$ *identically*, that is, no matter what values in A are given to the variables occurring in σ and τ. To make this precise, we need to analyze what is meant by the value in $\mathfrak{A}$ of an arbitrary term σ when values are given to its variables. It will be convenient to give values to *all* variables of the grammar G by introducing the notion of an $\mathfrak{A}$-*assignment*: This is simply a mapping b of V (the set of all variables of G) into A (the set of elements of the structure $\mathfrak{A}$). Such an $\mathfrak{A}$-assignment b determines a value in A for each term σ — a value we denote by $V_b^{\mathfrak{A}}(\sigma)$. To see how to define the mapping $V_b^{\mathfrak{A}}$ of T (the set of all terms) into A, consider an example.

Let us revert to the structure $\mathfrak{N} = (N, +, 0)$ and consider the term $\sigma = v_3 +_T (v_1 +_T v_3)$. Suppose b is an $\mathfrak{N}$-assignment such that $b(v_1) = 2$ and $b(v_3) = 7$. Then we compute $V_b^{\mathfrak{N}}(\sigma) = 16$ as follows. We first find $V_b^{\mathfrak{N}}(v_1 +_T v_3) = V_b^{\mathfrak{N}}(v_1) + V_b^{\mathfrak{N}}(v_3) = b(v_1) + b(v_3) = 2 + 7 = 9$, and then we compute $v_b^{\mathfrak{N}}(\sigma) = V_b^{\mathfrak{N}}(v_3) + V_b^{\mathfrak{N}}(v_1 +_T v_3) = b(v_3) + 9 = 7 + 9 = 16$. In short, $V_b^{\mathfrak{N}}(\sigma)$ is calculated by a recursive process, in which we successively compute values $V_b(\tau)$ for all terms τ which occur as parts of σ. In case the part τ is a variable in V, then $V_b(\tau) = b(\tau)$; in case τ is the constant symbol $\mathbf{0}$, then $V_b(\tau) = 0$; in case τ is

a compound term $\pi +_T \rho$, then $V_b^{\mathfrak{N}}(\tau) = V_b^{\mathfrak{N}}(\pi) + V_b^{\mathfrak{N}}(\rho)$. These observations can be summed up by saying that $V_b^{\mathfrak{N}}$ is a *homomorphism* of the algebra of terms $\mathfrak{T} = (T, +_T, \mathbf{0})$ into the structure $\mathfrak{N} = (N, +, 0)$, which agrees with b on the subset $V \subseteq T$.

The general process for computing the value of an arbitrary term σ, with respect to an arbitrary structure $\mathfrak{A}$ is similar, and rests upon the following.

(3.1) FUNDAMENTAL HOMOMORPHISM THEOREM (FHT). *For any structure $\mathfrak{A} = (A, +_A, 0_A)$, and any map $b: V \to A$, there exists a unique homomorphism h of $\mathfrak{T} = (T, +_T, 0)$ into $\mathfrak{A}$ such that $h(v) = b(v)$ for all $v \in V$.*[9] *We denote this homomorphism by $V_b^{\mathfrak{A}}$.*

As indicated above, we use the value maps $V_b^{\mathfrak{A}}$ to define the truth value $W^{\mathfrak{A}}(\psi)$, for any identity $\psi = \sigma \equiv \tau$ of G, by specifying that $W^{\mathfrak{A}}(\sigma \equiv \tau) = \mathsf{T}$ iff $V_b^{\mathfrak{A}}(\sigma) = V_b^{\mathfrak{A}}(\tau)$ for *every* $\mathfrak{A}$-assignment b.

In terms of these definitions we can now make precise the question that we put to our high school student in the introductory paragraphs of this paper.

(3.2) *Query*: Can we find an identity ψ of the grammar G such that $W^{\mathfrak{N}}(\psi) = \mathsf{T}$, but for which we do *not* have $\Gamma_{CAZ} \vdash \psi$?

In the next two sections we shall develop the machinery for demonstrating that this query has a negative answer. But first let us conclude the present section by two remarks about the Fundamental Homomorphism Theorem (FHT) for the algebra of terms $\mathfrak{T}$.

We shall not present a proof of the FHT here. The possibility of computing the unique values $h(\sigma)$ for each term σ, by recursion, should be intuitively clear from examples, and the main emphasis of the present paper lies in another direction. For those who would like to construct a proof for themselves, we remark that the definition of h by recursion over terms is very similar to the definition of numerical functions by recursion over the natural numbers. Thus, the FHT can be proved very much like the theorem justifying definition by recursion in the theory of natural numbers. There are two dual methods of proving the latter theorem,[10] and each method can be adapted to obtain a proof of the FHT.

The FHT can also be used to make precise the notion of substitution which appears in our rule of inference (R4) of Section 2. For let f be any mapping from V to T, and let Sub_f be the operation on T which maps each term σ to the term $\mathrm{Sub}_f(\sigma)$ obtained by substituting the term $f(v)$ throughout σ for each variable v occurring in σ. Clearly $\mathrm{Sub}_f(v) = f(v)$ for each $v \in V$, $\mathrm{Sub}_f(\mathbf{0}) = \mathbf{0}$, and $\mathrm{Sub}_f(\sigma +_T \tau) = \mathrm{Sub}_f(\sigma) +_T \mathrm{Sub}_f(\tau)$ for each $\sigma, \tau \in T$. By FHT we conclude that the operation Sub_f is nothing other than $V_f^{\mathfrak{T}}$ — and can be so defined. This observation leads to the following

(3.3) THEOREM ON SUBSTITUTION OF VALUES. *Let $\mathfrak{A} = (A, +_A, 0_A)$ be any structure. Let f be any $\mathfrak{T}$-assignment, b any $\mathfrak{A}$-assignment, σ any term. Then $V_b^{\mathfrak{A}}(\mathrm{Sub}_f \sigma) = V_{b'}^{\mathfrak{A}}(\sigma)$, where $b'(v) = V_b^{\mathfrak{A}}(f(v))$ for each $v \in V$.*

By way of proof, it suffices to note that the composite map $(V_b^{\mathfrak{A}} \circ \mathrm{Sub}_f)$ is a homomorphism of $\mathfrak{T}$ into $\mathfrak{A}$, and agrees with b' on V; the theorem then follows at once from the uniqueness part of FHT (3.1).

---

[9] A homomorphism h of $\mathfrak{T}$ into $\mathfrak{A}$ is a mapping $h: T \to A$ such that $h(\mathbf{0}) = 0_A$ and, for all $\sigma, \tau \in T$, $h(\sigma +_T \tau) = (h\sigma) +_A (h\tau)$. The existence of such a homomorphism for every structure $\mathfrak{A}$ and every map b, as in 3.1, is expressed by saying that the structure $\mathfrak{T}$ is *free* for the class of all structures $\mathfrak{A}$ (or simply that $\mathfrak{T}$ is *absolutely free*); and the set V, the domain of b, is called a *set of free generators* of $\mathfrak{T}$.

[10] See [3], where proofs of the theorem justifying definition by recursion are given for the system $\mathfrak{N}_S = (N, S, 0)$ of natural numbers with successor function S as fundamental operation. To adapt these proofs to the FHT for the system $\mathfrak{T} = (T, +_T, 0)$, one considers the operation S of $\mathfrak{N}_S$ to be analogous to $+_T$ in $\mathfrak{T}$, while 0 of $\mathfrak{N}_S$ is considered analogous to $\{0\} \cup V$ in $\mathfrak{T}$. In considering this analogy, the Peano axioms for $\mathfrak{N}_S$ should be compared with the conditions (1.2)–(1.4) for the system $\mathfrak{T}$. Obviously (1.2) is analogous to the Peano axiom $0 \neq Sx$ for all $x \in N$; (1.3) is analogous to the Peano axiom that $Sx = Sy$ implies $x = y$; and (1.4) is analogous to the induction axiom for $\mathfrak{N}_S$.

4. Soundness and completeness. As indicated in Section 2, the deductive system $(G, \vdash)$ can be considered sound only if it never leads us from true premises to a false conclusion. Now that truth values have been associated with the identities of G in a precise way (Section 3), we are in a position to prove soundness. It is convenient to formulate this in terms of a semantical notion of implication.

Let Γ be any set of identities of G and let ψ be any particular identity. A structure $\mathfrak{A}$ is said to *satisfy* ψ (resp. Γ) iff $W^{\mathfrak{A}}(\psi) = \mathsf{T}$ (resp. $W^{\mathfrak{A}}(\theta) = \mathsf{T}$ for every $\theta \in \Gamma$). We say that Γ *implies* ψ, or that ψ is a *consequence* of Γ, in case ψ is satisfied by every structure $\mathfrak{A}$ which satisfies Γ. We use the notation $\Gamma \models \psi$ to indicate that Γ implies ψ.

(4.1) THEOREM. *Whenever* Γ *is a set of identities and* ψ *is an identity of* G *such that* $\Gamma \vdash \psi$, *then also* $\Gamma \models \psi$.

It will be seen that this theorem simply expresses the proposition that the deductive system $(G, \vdash)$ never leads from true premises to a false conclusion. Accordingly, we call this the *Soundness Theorem* for the system $(G, \vdash)$.

The *proof* of this theorem is sufficiently straightforward so that most of the details can be omitted. Assuming that we have Γ and ψ such that $\Gamma \vdash \psi$, there must be (by definition of $\vdash$) a deduction, $(\theta_1, \ldots, \theta_n)$, from Γ, such that $\theta_n = \psi$. We now consider each θ_j in this deduction and we show (by induction on j) that θ is satisfied by every structure $\mathfrak{A}$ which satisfies Γ. To prepare the ground for this inductive proof it is convenient to establish five lemmas to the effect that $W^{\mathfrak{A}}(\theta) = \mathsf{T}$ for *all* $\mathfrak{A}$ in case θ is a tautology, and that if θ is inferred from one or two hypotheses by one of the rules of inference (R1)–(R4), then $W^{\mathfrak{A}}(\theta) = \mathsf{T}$ for every structure $\mathfrak{A}$ which satisfies these hypotheses. These lemmas are all established trivially, except perhaps for the one involving (R4), the rule of substitution — but this follows easily from the *Theorem on substitution of values* (3.3) of the preceding section. When the induction on j is completed, we obtain (4.1) by taking the case $j = n$.

The problem of establishing *completeness* for the deductive system $(G, \vdash)$ is not quite so simple as that of establishing its soundness. Here we must show that given Γ and ψ, it is always possible to find a deduction $(\theta_1, \ldots, \theta_n)$ from Γ, such that $\theta_n = \psi$, *except* when there is a structure $\mathfrak{A}$ which satisfies Γ but not ψ, i.e., except when we have *not* $\Gamma \models \psi$.

(4.2) THEOREM. (*Completeness.*) *Whenever* $\Gamma \models \psi$, *then also* $\Gamma \vdash \psi$.

Proof. Consider the contrapositive of this theorem: If *not* $\Gamma \vdash \psi$, then *not* $\Gamma \models \psi$. To prove this, we assume that Γ and ψ are such that *not* $\Gamma \vdash \psi$, and we seek to find $\mathfrak{A}$ such that $W^{\mathfrak{A}}(\phi) = \mathsf{T}$ for all $\phi \in \Gamma$, while $W^{\mathfrak{A}}(\psi) = \mathsf{F}$. It turns out that we can construct such an $\mathfrak{A}$ by using the given Γ, *but independent of* ψ. Thus the completeness theorem follows at once from the following stronger result.

(4.3) LEMMA. *For any set* Γ *of identities of* G *there is a structure* $\mathfrak{A}$, *such that for each* ψ *of* G *we have* $\mathfrak{A}$ *satisfies* ψ *iff* $\Gamma \vdash \psi$. *In particular,* $W^{\mathfrak{A}}(\psi) = \mathsf{F}$ *whenever* **not** $\Gamma \vdash \psi$.[11]

To construct $\mathfrak{A}$, we start from the algebra of terms $\mathfrak{T} = (T, +_T, \mathbf{0})$ of (1.1) and consider the relation I_Γ on T such that, for all $\sigma, \tau \in T$ we have

$$\sigma I_\Gamma \tau \text{ iff } \Gamma \vdash \sigma \equiv \tau.$$

This relation I_Γ is reflexive because $\Gamma \vdash \sigma \equiv \sigma$ for any $\sigma \in T$ by (R0); I_Γ is symmetric by (R1); and it is transitive by (R2). In short, I_Γ is an equivalence relation on T, and we can take A to be the partition of T induced by I_Γ. Thus, for every $\sigma \in T$ there is an element $[\sigma]$ of A, namely, $[\sigma] = \{\tau \in T \mid \sigma I_\Gamma \tau\}$. In particular, we take 0_A to be $[\mathbf{0}]$. In order to complete our description of the structure $\mathfrak{A}$, it remains

[11] Think of Γ as a set of axioms — e.g., axioms for the theory of groups. The lemma asserts that there is a structure $\mathfrak{R}$ satisfying these axioms, such that the only identities holding in $\mathfrak{A}$ are those derivable from the axioms. If indeed Γ is a set of axioms for group theory, then $\mathfrak{A}$ will be a *free group*. In general, we may call $\mathfrak{A}$ a *free* Γ-*structure*.

only to define the operation $+_A$ on A. This we do by stipulating that, for any $\sigma, \tau \in T$,

$$[\sigma]+_A[\tau] = [\sigma +_T \tau],$$

a definition justified by the fact that if $[\sigma] = [\sigma']$ and $[\tau] = [\tau']$, then $\sigma I_\Gamma \sigma'$ and $\tau I_\Gamma \tau'$, which means that we will have $(\sigma +_T \tau) I_\Gamma (\sigma' +_T \tau')$ by (R3), and thus $[\sigma +_T \tau] = [\sigma' +_T \tau']$.

Now that the structure $\mathfrak{A} = (A, +_A, 0_A)$ is defined, we wish to show that it satisfies the condition of Lemma 4.3. To do this, we need to have a way of evaluating $V_b^{\mathfrak{A}}(\sigma)$ for any term σ and any $\mathfrak{A}$-assignment b. This can be accomplished if, with each map $f: V \to T$, we associate the $\mathfrak{A}$-assignment b_f such that $b_f(v) = [f(v)]$ for every variable $v \in V$. (Obviously any $\mathfrak{A}$-assignment b will be equal to such a map b_f, for some f; we simply choose $f(v) \in b(v)$ for all $v \in V$.) Now we make the following

Claim: For every map $f: V \to T$, and every term $\sigma \in T$, we have

$$V_{b_f}^{\mathfrak{A}}(\sigma) = [\mathrm{Sub}_f \sigma].$$

To establish this claim, we consider the subset $U \subseteq T$ consisting of those terms σ such that $V_{b_f}^{\mathfrak{A}}(\sigma) = [\mathrm{Sub}_f \sigma]$ for every map $f: V \to T$. Certainly $0 \in U$ since

$$\begin{aligned} V_{b_f}^{\mathfrak{A}}(0) &= 0_A && \text{by definition of } V_{b_f}^{\mathfrak{A}} \\ &= [\mathbf{0}] && \text{by definition of } 0_A \\ &= [\mathrm{Sub}_f\, \mathbf{0}] && \text{by definition of } \mathrm{Sub}_f. \end{aligned}$$

Similarly, for every variable $v \in V$ we have $v \in U$, since

$$\begin{aligned} V_{b_f}^{\mathfrak{A}}(v) &= b_f(v) && \text{by definition of } V_{b_f}^{\mathfrak{A}} \\ &= [f(v)] && \text{by definition of } b_f \\ &= [\mathrm{Sub}_f(v)] && \text{by definition of } \mathrm{Sub}_f. \end{aligned}$$

Finally, U is closed under $+_T$. For if $\sigma, \tau \in U$ we have

$$\begin{aligned} V_{b_f}^{\mathfrak{A}}(\sigma +_T \tau) &= V_{b_f}^{\mathfrak{A}}(\sigma) +_A V_{b_f}^{\mathfrak{A}}(\tau) && \text{by definition of } V_{b_f}^{\mathfrak{A}}, \\ &= [\mathrm{Sub}_f(\sigma)] +_A [\mathrm{Sub}_f(\tau)] && \text{since } \sigma, \tau \in U, \\ &= [\mathrm{Sub}_f(\sigma) +_T \mathrm{Sub}_f(\tau)] && \text{by definition of } +_A, \\ &= [\mathrm{Sub}_f(\sigma +_T \tau)] && \text{by definition of } \mathrm{Sub}_f, \end{aligned}$$

so that $(\sigma +_T \tau) \in U$ by definition of U.

Having shown that $0 \in U$, $V \subseteq U$, and U is closed under $+_T$, it follows by (1.4) that $U = T$, i.e., that $V_{b_f}^{\mathfrak{A}}(\sigma) = [\mathrm{Sub}_f \sigma]$ for *every* term σ and every map $f: V \to T$. Thus our *Claim* is established.

Now consider the injection map $f^*: V \to T$ such that $f^*(v) = v$ for all $v \in V$. Clearly Sub_{f^*} is the identity map on T, by FHT (3.1), since the identity map is certainly a homomorphism of $\mathfrak{T}$ into itself and it agrees with f^* on V. Hence, letting b^* be the $\mathfrak{A}$-assignment such that $b^*(v) = b_{f^*}(v) = [v]$ for each $v \in V$, we obtain as a special case of our *Claim* that

(4.4) $$V_{b^*}^{\mathfrak{A}}(\sigma) = [\sigma] \text{ for all } \sigma \in T.$$

Also, if b is an arbitrary $\mathfrak{A}$-assignment, we can choose a map $f: V \to T$ so that

(4.5) $$V_b^{\mathfrak{A}}(\rho) = [\mathrm{Sub}_f \rho] \text{ for all } \rho \in T,$$

according to our *Claim*, for we have seen that we obtain $b_f = b$ by taking $f(v) \in b(v)$ for all v.

The machinery is now set for proving Lemma 4.3. Consider, first, an arbitrary identity ψ of the

grammar G — say ψ is $\sigma \equiv \tau$ for some $\sigma, \tau \in T$. In case *not* $\Gamma \vdash \psi$, we have *not* $\sigma I_\Gamma \tau$ (by definition of I_Γ), so that $[\sigma] \neq [\tau]$. Hence by (4.4) we get $V_{b^*}^{\mathfrak{A}}(\sigma) \neq V_{b^*}^{\mathfrak{A}}(\tau)$, which shows that $W^{\mathfrak{A}}(\phi) = \mathsf{F}$ by definition of $W^{\mathfrak{A}}$.

On the other hand consider the case $\Gamma \vdash \psi$, where ψ is $\sigma \equiv \tau$, and take an arbitrary $\mathfrak{A}$-assignment b. Using (4.5), let us choose a map $f: V \to T$ such that $V_b^{\mathfrak{A}}(\rho) = [\mathrm{Sub}_f \rho]$ for all $\rho \in T$, and then, with this f, let us apply (R4) to obtain $\Gamma \vdash (\mathrm{Sub}_f \sigma) \equiv (\mathrm{Sub}_f \tau)$. This last relation gives $(\mathrm{Sub}_f \sigma) I_\Gamma (\mathrm{Sub}_f \tau)$ by definition of I_Γ, so that $[\mathrm{Sub}_f \sigma] = [\mathrm{Sub}_f \tau]$. But then $V_b^{\mathfrak{A}}(\sigma) = V_b^{\mathfrak{A}}(\tau)$ by choice of f. Since b was an *arbitrary* $\mathfrak{A}$-assignment, this shows that $W^{\mathfrak{A}}(\sigma \equiv \tau) = \mathsf{T}$, i.e., $W^{\mathfrak{A}}(\psi) = \mathsf{T}$, completing the proof of Lemma (4.3).

As indicated above, the completeness (4.2) of the deductive system $(G, \vdash)$ is an immediate corollary of Lemma 4.3.

5. Structural operations preserving identities. The notion of derivability, symbolized $\Gamma \vdash \phi$, is defined in terms of rules of inference whose applicability depends only on the arrangement of symbols in formulas, not on any meanings of these formulas. For this reason we call derivability a *syntactical* notion. By contrast, the notion of implication, symbolized $\Gamma \vDash \phi$, is defined in terms of the values $V_b^{\mathfrak{A}}(\sigma)$ of terms σ and of the truth or falsity $W^{\mathfrak{A}}(\psi)$ of formulas ψ, which depend on meanings, and so implication is called a *semantical* notion. By the soundness and completeness results of the preceding section, we have $\Gamma \vdash \phi$ if and only if $\Gamma \vDash \phi$, for all equational identities ϕ and all sets Γ of such identities.

Now the question raised in the opening paragraphs of this paper and made precise in Query (3.2) involves both a semantical notion, $W^{\mathfrak{N}}$, and a syntactical notion, $\vdash$. However, using the soundness and completeness theorems we see that the question is equivalent to one formulated purely in semantical terms: Is there any equational identity ψ of the grammar G such that $W^{\mathfrak{N}}(\psi) = \mathsf{T}$, but *not* $\Gamma_{\mathrm{CAZ}} \vDash \psi$?

Recall that the relation $\Gamma_{\mathrm{CAZ}} \vDash \psi$ will hold just in case ψ is true in every structure $\mathfrak{A}$ which satisfies the CAZ laws. Now such a structure $\mathfrak{A}$ is called a *commutative semigroup with identity element.* Thus our original question becomes transformed to the following.

(5.0) *Query*: Can we find an equational identity ψ of the grammar G such that $W^{\mathfrak{N}}(\psi) = \mathsf{T}$, but such that $W^{\mathfrak{A}}(\psi) = \mathsf{F}$ for at least one commutative semigroup with identity element, $\mathfrak{A}$?

In this section we shall demonstrate that the answer to this question is negative. We shall show that whenever an identity ψ holds in $\mathfrak{N}$ it must also hold in every commutative semigroup with identity, $\mathfrak{A}$. We shall do this by indicating how $\mathfrak{A}$ can be obtained from $\mathfrak{N}$ by means of the structural operations of forming subalgebras, homomorphic images, and direct products. These operations are widely known in the context of groups and rings, but in fact they are applicable to algebraic structures of completely arbitrary type. We shall briefly review the definition of each of these operations, below, indicating how each of them may be shown to preserve the truth of equational identities.

(5.1) DEFINITION. *Let* $\mathfrak{A}_1$ *and* $\mathfrak{A}_2$ *be two structures,* $\mathfrak{A}_i = (A_i, +_i, 0_i)$ *for* $i = 1, 2$. *We say that* $\mathfrak{A}_1$ *is a substructure of* $\mathfrak{A}_2$ *iff*: (1) $A_1 \subseteq A_2$, (2) $x +_2 y = x +_1 y$ *for all* $x, y \in A_1$, *and* (3) $0_1 = 0_2$.[12]

(5.2) THEOREM. *If* $\mathfrak{A}_1$ *is a substructure of* $\mathfrak{A}_2$, *if* ψ *is an equational identity of the grammar G, and if* $W^{\mathfrak{A}_2}(\psi) = \mathsf{T}$, *then also* $W^{\mathfrak{A}_1}(\psi) = \mathsf{T}$.

The intuitive content of this theorem is obvious: If an equation holds whenever values in A_2 are assigned to its variables, and if $A_1 \subseteq A_2$, then of course the equation holds whenever values in A_1 are assigned to its variables. Using our technical definition of $W^{\mathfrak{A}_1}$, we prove the theorem by first showing

[12] Condition (2) can be expressed by saying that $+_1$ is the *restriction* of $+_2$ to A_1. Sometimes the term *subalgebra* is used instead of *substructure.* Note that if I is the set of integers and $+'$ the usual operation of addition on it, then the structure $(I, +', 0)$ has $\mathfrak{N}$ as a substructure. However the former is a group, $\mathfrak{N}$ is not (as elements other than 0 have no inverses), so a group may have a substructure which is not a subgroup.

that for any assignment $b : V \to A_1$ we have $V_b^{\mathfrak{A}_1}(\sigma) = V_b^{\mathfrak{A}_2}(\sigma)$ for every term σ. This, in turn, is accomplished by employing conditions (1)–(3) of (5.1) and making a suitable application of (1.4).

(5.3) DEFINITION. *Let $\mathfrak{A}_1$ and $\mathfrak{A}_2$ be two structures, $\mathfrak{A}_i = (A_i, +_i, 0_i)$. By a homomorphism of $\mathfrak{A}_1$ onto $\mathfrak{A}_2$ is meant a surjective mapping $h : A_1 \to A_2$ such that*: (1) $h(x +_1 y) = h(x) +_2 h(y)$ *for all* $x, y \in A_1$, *and* (2) $h(0_1) = 0_2$. *We say that $\mathfrak{A}_1$ is* ***homomorphic*** *to $\mathfrak{A}_2$, or that $\mathfrak{A}_2$ is a* ***homomorphic image*** *of $\mathfrak{A}_1$, in case there exists a homomorphism of $\mathfrak{A}_1$ onto $\mathfrak{A}_2$.*[13]

(5.4) THEOREM. *If $\mathfrak{A}_1$ is homomorphic to $\mathfrak{A}_2$, if ψ is an equational identity of the grammar G, and if $W^{\mathfrak{A}_1}(\psi) = \mathsf{T}$, then also $W^{\mathfrak{A}_2}(\psi) = \mathsf{T}$.*

Although not quite as obvious as (5.2), the content of this theorem is familiar, at least in an intuitive way, to most mathematicians. To obtain a proof of (5.4) using our technical definition of $W^{\mathfrak{A}_i}$ again requires us first to establish an appropriate lemma involving the values $V_{b'}^{\mathfrak{A}_1}(\sigma)$ for arbitrary terms σ of the grammar G. In this case we must first show that for each assignment $b : V \to A_2$ we can find an assignment $b' : V \to A_1$ such that $h(b'(v)) = b(v)$ for every variable $v \in V$, which we can do because the homomorphism $h : A_1 \to A_2$ is surjective. And then we show that for every term $\sigma \in T$ we have $h(V_{b'}^{\mathfrak{A}_1}(\sigma)) = V_b^{\mathfrak{A}_2}(\sigma)$ by applying (1.4) and using conditions (1) and (2) of (5.3).

(5.5) DEFINITION. *Let J be a non-empty set (whose elements are called indexes), and for each $j \in J$ let $\mathfrak{A}_j$ be a structure, $\mathfrak{A}_j = (A_j, +_j, 0_j)$. By the* ***direct product*** *$P_{j\in J}\,\mathfrak{A}_j$ of the structures $\mathfrak{A}_j$ we mean the structure $\mathfrak{B} = (B, +, 0)$ obtained as follows: B is the Cartesian product $P_{j\in J} A_j$ of the sets A_j, i.e., B is the set of all those mappings $f : J \to \bigcup_{j\in J} A_j$ such that $f(j) \in A_j$ for every $j \in J$; $+$ is the operation on B such that, for any $f, g \in B$, we have $(f + g)(j) = f(j) +_j g(j)$ for every $j \in J$; and 0 is the element of B such that $0(j) = 0_j$ for each $j \in J$. In case there is an $\mathfrak{A}$ such that $\mathfrak{A}_j = \mathfrak{A}$ for every $j \in J$, we write ${}^J\mathfrak{A}$ for $P_{j\in J}\,\mathfrak{A}_j$, and we call ${}^J\mathfrak{A}$ the* ***J-direct power*** *of $\mathfrak{A}$.*

(5.6) THEOREM. *Suppose that $\mathfrak{B} = P_{j\in J}\,\mathfrak{A}_j$ as in (5.5). If ψ is any equational identity of the grammar G such that $W^{\mathfrak{A}_j}(\psi) = \mathsf{T}$ for every $j \in J$, then also $W^{\mathfrak{B}}(\psi) = \mathsf{T}$.*[14] *In particular, if $\mathfrak{B}$ is the J-direct power of $\mathfrak{A}$, and if $W^{\mathfrak{A}}(\psi) = \mathsf{T}$, then also $W^{\mathfrak{B}}(\psi) = \mathsf{T}$.*

As in Theorems (5.2) and (5.4), the proof of (5.6) depends on establishing a suitable lemma concerning the values of an arbitrary term in the several structures under consideration. In the present case, to every assignment $b : V \to B$ let us associate the assignments $b_j : V \to A_j$, for each $j \in J$, by the rule that $b_j(v) = (b(v))(j)$ for all $v \in V$. Then for any term $\sigma \in T$ we can show that $(V_b^{\mathfrak{B}}(\sigma))(j) = V_{b_j}^{\mathfrak{A}_j}(\sigma)$ for every $j \in J$. This lemma requires for its proof an application of (1.4) and Definition (5.5); Theorem (5.6) follows at once from the lemma and the definition of $W^{\mathfrak{B}}$.

All of the necessary machinery is now at hand to resolve the query in the opening paragraphs of the paper.

(5.7) THEOREM. *Let ψ be any equational identity of the grammar G, and let $\mathfrak{A}$ be any commutative semigroup with identity element (i.e., any structure satisfying the CAZ identities ϕ_C, ϕ_A, and ϕ_Z of (2.1)). If $W^{\mathfrak{N}}(\psi) = \mathsf{T}$ then also $W^{\mathfrak{A}}(\psi) = \mathsf{T}$.*

Proof. We shall show how to construct structures $\mathfrak{B}_1$ and $\mathfrak{B}_2$ such that $\mathfrak{B}_1$ is a J-direct power of $\mathfrak{N}$ for some index set J, $\mathfrak{B}_2$ is a substructure of $\mathfrak{B}_1$, and $\mathfrak{A}$ is a homomorphic image of $\mathfrak{B}_2$. Using the hypothesis $W^{\mathfrak{N}}(\psi) = \mathsf{T}$ we will then be able to conclude $W^{\mathfrak{B}_1}(\psi) = \mathsf{T}$ by (5.6), $W^{\mathfrak{B}_2}(\psi) = \mathsf{T}$ by (5.2), and finally $W^{\mathfrak{A}}(\psi) = \mathsf{T}$ by (5.4), thus proving (5.7).

[13] A homomorphism which is injective (one-one) is called an *isomorphism*; $\mathfrak{A}_1$ and $\mathfrak{A}_2$ are said to be *isomorphic* if there is an isomorphism from one onto the other (in which case there are always isomorphisms in each direction).

[14] The converse of (5.6) also holds, unlike the situations for theorems (5.2) and (5.4). In fact, the converse of (5.6) may be obtained by applying (5.4), since $P_{j\in J}\,\mathfrak{A}_j$ is homomorphic to $\mathfrak{A}_j$ for every $j \in J$. To obtain a homomorphism h_j of $P_{j\in J}\,\mathfrak{A}_j$ onto $\mathfrak{A}_j$, we set $h_j(f) = f(j)$ for all $f \in P_{j\in J} A_j$.

In fact we take $\mathfrak{B}_1 = {}^A\mathfrak{N}$, using the universe A of the given structure $\mathfrak{A}$ as the index set for forming a direct power of $\mathfrak{N}$. Say $\mathfrak{B}_1 = (B_1, +_1, 0_1)$.

Now let B_2 be the subset of B_1 such that $f \in B_2$ if and only if $f(a) = 0$ for all but finitely many $a \in A$. It is evident that $0_1 \in B_2$, and that B_2 is closed under $+_1$. Hence, letting $+_2$ be the restriction of $+_1$ to B_2, putting $0_2 = 0_1$, and taking $\mathfrak{B}_2 = (B_2, +_2, 0_2)$, we see that $\mathfrak{B}_2$ is a substructure of $\mathfrak{B}_1$.

Finally, we wish to show that $\mathfrak{B}_2$ is homomorphic to $\mathfrak{A}$. Recall that for $f \in B_2$ and $a \in A$ we have $f(a) \in N$; thus $f(a) \cdot a$ will denote, as usual, the result of adding the element a to itself $f(a)$ times (in the structure $\mathfrak{A}$). Since $f(a) = 0$ for all but finitely many $a \in A$, we can form the sum $\Sigma_{a \in A}(f(a) \cdot a)$ in the structure $\mathfrak{A}$, for each $f \in B_2$. Now setting $h(f) = \Sigma_{a \in A}(f(a) \cdot a)$ for $f \in B_2$, we see that we have a mapping $h: B_2 \to A$; we claim that h is the required homomorphism of $\mathfrak{B}_2$ onto $\mathfrak{A}$.

Indeed, h is certainly surjective, for given any $a \in A$ we may consider the $f \in B_2$ such that $f(a) = 1$ and $f(x) = 0$ for every $x \in A \sim \{a\}$, and we easily calculate $h(f) = a$. Furthermore $h(0_2) = 0$, since 0_2 is the element of B_2 such that $0_2(a) = 0$ for all $a \in A$. Thus it remains only to show that $h(f +_2 g) = h(f) +_A h(g)$ for all $f, g \in B_2$, which can be done by a simple computation that we leave to the reader. The proof of (5.7) is thus completed.

(5.8) Corollary. *If* $W^{\mathfrak{N}}(\psi) = \mathsf{T}$ *then* $\{\phi_C, \phi_A, \phi_Z\} \models \psi$.

This is immediate from (5.7) by the definition of $\models$.

(5.9) Corollary. *If* $W^{\mathfrak{N}}(\psi) = \mathsf{T}$ *then* $\{\phi_C, \phi_A, \phi_Z\} \vdash \psi$.

This is immediate from (5.8) by the Completeness Theorem (4.1).

Corollary (5.9) asserts that any equational identity which holds in $\mathfrak{N}$ must be derivable from the CAZ identities using tautologies (R0) and the rules of inference (R1)–(R4). Our original query is settled.

Since $W^{\mathfrak{N}}(\psi) = T$ for $\psi \in \{\phi_C, \phi_A, \phi_Z\}$ — i.e., the CAZ laws are true for $\mathfrak{N}$ — the converse of 5.8 holds by definition of $\models$, and hence the converse of 5.9 holds by the Soundness Theorem (4.1).

Now whenever we have a structure $\mathfrak{A}$, and a set Γ of equational identities such that $W^{\mathfrak{A}}(\psi) = T$ iff $\Gamma \vdash \psi$ for all identities ψ, we call Γ a *base* for the equational theory of $\mathfrak{A}$. Thus 5.9 together with its converse are expressed in the following.

(5.10) Theorem. $\{\phi_C, \phi_A, \phi_Z\}$ *is a base for the equational theory of* $\mathfrak{N}$.

6. Identities in structures other than $\mathfrak{N}$. In this section we describe informally, and without proofs, the extension of Theorem 5.10 to certain other structures, closely related to $\mathfrak{N}$, which come naturally to mind. And we end with an open problem.

Let us start with the structure $\mathfrak{I} = (I, +, 0)$ — the additive group of integers. Of course each of the CAZ laws, and hence any identity derivable from them, holds in $\mathfrak{I}$. But conversely, if ψ is any identity holding in $\mathfrak{I}$, then it holds in the substructure $\mathfrak{N}$ of $\mathfrak{I}$ (by 5.2), and so ψ is derivable from the CAZ laws (by 5.10). Thus we see that $\{\phi_C, \phi_A, \phi_Z\}$ *is a base for the equational theory of* $\mathfrak{I}$.

Consider next the structure $\mathfrak{P} = (P, \cdot, 1)$ — the multiplicative semigroup (with identity) of positive integers. $\mathfrak{P}$ has a substructure (consisting of the powers of 2) that is mapped homomorphically by $\log_2$ onto $\mathfrak{N}$, from which we infer (by 5.2 and 5.4) that any identity holding in $\mathfrak{P}$ must also hold in $\mathfrak{N}$ and hence (by 5.10) is derivable from the CAZ laws. It follows that $\{\phi_C, \phi_A, \phi_Z\}$ *is a base for the equational theory of* $\mathfrak{P}$.[15]

Since $\mathfrak{P}$ is a substructure of the multiplicative semi-groups $\mathfrak{N}' = (N, \cdot, 1)$ and $\mathfrak{I}' = (I, \cdot, 1)$, the last result can be combined with 5.2 to conclude that $\{\phi_C, \phi_A, \phi_Z\}$ *is a base for the equational theories of* $\mathfrak{N}'$ *and of* $\mathfrak{I}'$. However, there is one point about this assertion that may, at first, appear suspicious. Namely, 0 is an element of N and of I, and in both $\mathfrak{N}'$ and $\mathfrak{I}'$ we have $x \cdot 0 = 0$ for all elements x of the structure. Doesn't this give us an identity not derivable from the CAZ laws?

---

[15] Note that when interpreted with respect to P, the "zero law" Z expresses the fact that $x \cdot 1 = x$ for all $x \in P$.

Paradoxically, the answer to this question is "yes and no." The point is that the grammar G possesses only one individual constant, the symbol $\mathbf{0}$, and when G is interpreted with respect to either of the structures $\mathfrak{N}'$ or $\mathfrak{I}'$, this symbol denotes the unique distinguished element of the structure, namely, the number 1. Thus there is no symbol of G denoting the number 0, and so the fact that the equation $x \cdot 0 = 0$ holds identically in $\mathfrak{N}'$ and in $\mathfrak{I}'$, cannot be expressed by a formula of the grammar G. The only formulas of G which hold in $\mathfrak{N}'$ or in $\mathfrak{I}'$ are those derivable from the CAZ laws.

Of course if we pass from $\mathfrak{I}'$ to the system $(I, \cdot, 1, 0)$, with two distinguished elements, then we will need a richer grammar than G to express its equational identities. Namely, we choose a grammar with an algebra of terms $(T, +_T, \mathbf{1}, \mathbf{0})$ possessing two individual constants, and within this grammar we can choose a base for the equational theory of $(I, \cdot, 1, 0)$ which includes both of the identities $(v_0 +_T \mathbf{1}) \equiv v_0$ and $(v_0 +_T \mathbf{0}) \equiv \mathbf{0}$.

A more interesting structure is the ring $\mathfrak{I}'' = (I, +, \cdot, 0, 1)$, possessing two binary operations as well as two distinguished elements. Using an equational grammar G'' with two operation symbols and two individual constants, we can show that *the following identities constitute a base for the equational theory of* $\mathfrak{I}''$: The CAZ laws for each of the reduced systems $(I, +, 0)$ and $(I, \cdot, 1)$, the distributive law for $\cdot$ over $+$, and the law $v_0 \cdot_T \mathbf{0} \equiv \mathbf{0}$.

Let Γ be the set of the preceding 8 identities. While it is possible to establish that Γ is a base for the equational theory of $\mathfrak{I}''$ along the lines we employed in Section 5 to obtain a base for the theory of $\mathfrak{N}$, there is a very different method which will seem natural to anyone who has studied the elements of ring theory.

Namely, we consider the set T'' of all terms of the grammar G'', and we single out a subset F of T'' whose elements we call *formal polynomials.* We shall not give a precise definition of F, but simply indicate that we first define *numerals* as being terms of the form 0 or $1 +_T 1 +_T \cdots +_T 1$ (suitably parenthesized), we define formal powers of a variable v_i as terms of the form $v_i \cdot_T v_i \cdot_T \cdots \cdot_T v_i$, and we define formal *monomials* as formal products of a formal numeral followed by a string of formal powers of variables in alphabetical order. Finally, the formal polynomials are defined as formal sums of a string of formal monomials arranged in some prescribed, alphabetic-type order. The definition is such as to allow us to prove the following basic lemmas.

(6.1) LEMMA. *For any term σ of T'' there is a formal polynomial τ of T'' such that* $\Gamma \vdash \sigma \equiv \tau$.

(6.2) LEMMA *If τ and ρ are any formal polynomials of T'' such that $W^{\mathfrak{I}''}(\tau \equiv \rho) = \mathsf{T}$, then $\tau = \rho$.*

The proof of 6.1 provides an explicit recipe for constructing τ when σ is given, and for constructing a formal derivation of $\sigma \equiv \tau$ from Γ. The proof of 6.2 depends on the fact that a non-zero polynomial function (of one argument) has only finitely many roots over $\mathfrak{I}''$, while I has infinitely many elements. From these lemmas we easily obtain the desired result that Γ is a base for the equational theory of $\mathfrak{I}''$. Furthermore, the lemmas lead to a *decision procedure* for this equational theory, namely, an automatic process to decide, in a finite number of steps, whether or not a given identity $\sigma \equiv \tau$ holds for $\mathfrak{I}''$. Namely, being given any terms σ and τ we first find formal polynomials σ' and τ' such that $\Gamma \vdash \sigma \equiv \sigma'$ and $\Gamma \vdash \tau = \tau'$. Since $W^{\mathfrak{I}''}(\psi) = \mathsf{T}$ for each $\psi \in \Gamma$, this shows that $W^{\mathfrak{I}''}(\sigma \equiv \sigma') = \mathsf{T}$ and $W^{\mathfrak{I}''}(\tau \equiv \tau') = \mathsf{T}$. Hence by (6.2), we have $W^{\mathfrak{I}''}(\sigma \equiv \tau) = \mathsf{T}$ iff $\sigma' = \tau'$, which furnishes the required decision procedure.[16] Furthermore, if $W^{\mathfrak{I}''}(\sigma \equiv \tau) = \mathsf{T}$ we see that $\Gamma \vdash \sigma \equiv \sigma'$ and $\Gamma \vdash \tau \equiv \sigma'$, and hence $\Gamma \vdash \sigma \equiv \tau$ by definition of $\vdash$, which assures us that Γ is indeed a base for the equational theory of $\mathfrak{I}''$.

The same proof shows that Γ is a base for the equational theory of any infinite integral domain. And of course Γ is also a base for the equational theory of the semi-ring $\mathfrak{N}'' = (N, +, \cdot, 0, 1)$, by 5.2, since $\mathfrak{N}''$ is a substructure of $\mathfrak{I}''$.

---

[16] This is because there is an automatic method to decide, being given terms τ and τ', whether or not they are the same.

We come to an unsolved problem of equational logic if we enrich the structure $\mathfrak{N}''$ by means of the exponential operation exp on N, such that $x \exp y = x^y$ for all $x, y \in N$. Setting $\mathfrak{N}''' = (N, +, \cdot, \exp, 0, 1)$, and passing to a grammar based on 3 operation symbols and 2 individual constants, one may seek a base for the equational theory of $\mathfrak{N}'''$. In particular, Alfred Tarski[17] has raised the question whether we can obtain a base for this theory by starting from the base Γ for the theory of $\mathfrak{N}''$, and adjoining the following identities, almost as well known to high school students as those of Γ itself:

$$
(6.3)\qquad
\begin{array}{llll}
\text{(a)} & (v_0 \exp_T \mathbf{0}) \equiv \mathbf{1} & \text{(d)} & (v_0 \exp_T (v_1 +_T v_2)) \equiv ((v_0 \exp_T v_1) \cdot_T (v_0 \exp_T v_2)) \\
\text{(b)} & (v_0 \exp_T \mathbf{1}) \equiv v_0 & \text{(e)} & ((v_0 \cdot_T v_1) \exp_T v_2) \equiv ((v_0 \exp_T v_2) \cdot_T (v_1 \exp_T v_2)) \\
\text{(c)} & (\mathbf{1} \exp_T v_0) \equiv \mathbf{1} & \text{(f)} & (v_0 \exp_T (v_1 \cdot_T v_2)) \equiv ((v_0 \exp_T v_1) \exp_T v_2).
\end{array}
$$

Although mathematicians have been trying to answer this question for quite a few years, it remains unsolved.

However, if we reduce the structure $\mathfrak{N}'''$ to the structure $\mathfrak{N}^* = (N, +, \cdot, \exp)$ which has no distinguished elements, then the set of identities Δ consisting of Γ together with the identities (d)–(f) of 6.3 above, has been shown by Tarski's student Charles Martin *not* to be a base for the equational theory of $\mathfrak{N}^*$. Martin showed that the following identity, which obviously holds in $\mathfrak{N}^*$, is *not* derivable from Δ:

$$(x^y + x^y)^x \cdot (y^x + y^x)^y \equiv (x^x + x^x)^y \cdot (y^y + y^y)^x.$$

(We have used the usual exponential notation for clarity.) And indeed, he showed that the equational theory of $\mathfrak{N}^*$ *has no finite base.*[18]

7. Equational identities and ordinary equations. In Sections 2–6 we have developed the logic of equational identities. We have provided a syntactically-defined deductive system, as well as a semantical notion of consequence, for a grammar made up of these identities, and we have related these notions by theorems of soundness and completeness. We have shown that equational identities are preserved under formation of substructures, homomorphic images, and direct products, and we have used these facts to obtain bases for the equational theories of several familiar number systems.

In this final section of our paper we wish to return to the problem of those high school students, mentioned in our introductory remarks, who continue to struggle bravely with the problem of solving equations in their algebra courses. As we pointed out, the equations they seek to solve normally do *not* hold identically in some number system, but possess only one or two roots. What sort of logical systems underlie the solution of such equations?

To examine this question, let us simplify the details by returning to a consideration of the number system $\mathfrak{N} = (N, +, 0)$, with only one operation and one distinguished element. In order to formulate equations over $\mathfrak{N}$, we provide ourselves with a grammar G^* which has exactly the same algebra of terms $\mathfrak{T} = (T, +_T, \mathbf{0})$ as our original grammar G, but instead of the identity symbol $\equiv$ of G, the

[17] Alfred Tarski, along with Garrett Birkhoff, may be regarded as founders of the systematic study of equational logic. Birkhoff's early paper [1] contains the ideas which underlie the completeness theorem for equational logic, even though this is not explicitly formulated; Tarski's recent paper [7] contains a description of current activity in the many areas of equational logic. Tarski's formulation of the problem of finding a base for the equational theory of exponentiation, multiplication, and addition was actually given for the system of *positive* integers, 0 being excluded because of the ambiguity in a possible definition of 0^0. If we take $0^0 = 1$, as assumed in our text above, we get the identity $x^0 \equiv 1$ which we have taken as (a) above; if we take $0^0 = 0$, we lose that identity but obtain $0^x \equiv 0$ instead. Thus in Tarski's formulation the grammar contains no symbol $\mathbf{0}$, and our identity (a) is omitted from the conjectured base.

[18] Charles Martin was Tarski's doctoral student. In addition to the results mentioned above, his doctoral dissertation provides finite bases for the systems $(N, \exp)$ and $(N, \exp, \cdot)$. See [4].

grammar G^* will contain an equals symbol $=$, and its formulas (to be called equations) will consist of strings of symbols of the form $\sigma^\frown = {}^\frown\tau$, for any terms $\sigma, \tau \in T$.[19]

When the grammar G^* is interpreted with respect to $\mathfrak{N}$, or to any structure $\mathfrak{A} = (A, +_A, 0_A)$ of the same type, it is important to keep in mind that an equation of G^* does not assume a definite truth value, T or F. For example, the equation $(v_2 +_T \mathbf{0}) = (v_1 +_T v_3)$ will be true in $\mathfrak{N}$ if we assign values 0 to v_1, 6 to v_2, and 6 to v_3, but the same equation will be false in $\mathfrak{N}$ if we assign 2 to v_1, 3 to v_2, and 4 to v_3. Thus, in place of the notation $W^{\mathfrak{A}}(\psi)$ used in connection with the grammar G to indicate the truth value assumed by the identity ψ with respect to the structure $\mathfrak{A}$, we shall now be using the notation $W_b^{\mathfrak{A}}(\alpha)$ to indicate the truth value assumed by the equation α of G^*, with respect to the structure $\mathfrak{A}$, when the variables occurring in α are given values in $\mathfrak{A}$ by the assignment $b : V \to A$.

With this change, we now define a semantical relation of consequence $\models^*$, relating a set Δ of equations of G^* with a single equation α, as follows.

(7.1) Definition. *$\Delta \models^* \alpha$ iff, whenever $\mathfrak{A}$ and $b : V \to A$ are such that $W_b^{\mathfrak{A}}(\delta) = \mathsf{T}$ for every $\delta \in \Delta$, then also $W_b^{\mathfrak{A}}(\alpha) = \mathsf{T}$.*

To see the difference between the relations $\models$ and $\models^*$, observe that we have

(7.2)
$$\{(v_0 +_T v_1) \equiv (v_1 +_T v_0)\} \models (v_0 +_T v_2) \equiv (v_2 +_T v_0), \text{ but}$$
$$\textit{not } \{(v_0 +_T v_1) = (v_1 +_T v_0)\} \models^* (v_0 +_T v_2) = (v_2 +_T v_0).$$

The second of these assertions can be verified from (7.1) by choosing $\mathfrak{A}$ to be any non-commutative group, and then defining $b : V \to A$ so that the elements $b(v_0)$ and $b(v_1)$ commute (say $b(v_0) = b(v_1)$), while $b(v_0)$ and $b(v_2)$ do not.

Now how do we obtain a syntactically defined relation $\vdash^*$ of derivability for our grammar G^*? Very simply. Let us begin by considering each of the rules (R0)–(R4) of Section 2, and forming corresponding rules (R0)*–(R4)* for the grammar G^* by changing the symbol $\equiv$ to $=$ throughout each one. It is easy to check that $W_b^{\mathfrak{A}}(\alpha) = \mathsf{T}$ for each equation α of the form $\sigma = \sigma$ which is certified as an axiom by (R0)*, and similarly, if α is inferred from one or two premises by one of the rules (R1)*–(R3)*, then $W_b^{\mathfrak{A}}(\alpha) = \mathsf{T}$ for any $\mathfrak{A}$ and b such that $W_b^{\mathfrak{A}}(\beta) = \mathsf{T}$ for each premise β. On the other hand, it is quite possible for α to be inferred from β by (R4)*, and to find $\mathfrak{A}$ and b for which $W_b^{\mathfrak{A}}(\beta) = \mathsf{T}$ but $W_b^{\mathfrak{A}}(\alpha) = \mathsf{F}$ — as we have shown in the second part of (7.2) above.

With these observations in mind, let us define a *formal deduction from* Δ — where Δ is any set of equations of G^* — to be a finite sequence $(\alpha_0, \ldots, \alpha_n)$ of equations such that each α_j is either an element of Δ, an instance of (R0)*, or else is inferred by one of the rules (R1)*–(R3)* from one or two earlier equations of the sequence. And we define $\Delta \vdash^* \alpha$ to hold if, and only if, there is such a deduction from Δ whose last equation is α.

With this slight modification of the notion of derivability, we find no trouble in adapting the proofs of both the Soundness Theorem (4.1), and the Completeness Theorem (4.2), first developed for the grammar G, to the grammar G^*.

(7.3) Theorem. *For any set Δ of equations of the grammar G^*, and any such equation α, we have $\Delta \vdash^* \alpha$ iff $\Delta \models^* \alpha$.*

From this result, it appears that a sound and complete system of deduction in the grammar of equations is essentially obtained by simplification from the system of deduction of the grammar of equational identities — a simplification consisting simply in dropping one rule of inference, (R4), the rule of substitution. The resulting rules, (R0)*–(R3)*, are of so elementary and transparent a

[19] Henceforth we shall write simply $\sigma = \tau$ in place of $\sigma^\frown = {}^\frown\tau$. Of course the shape of a particular symbol is irrelevant to its mathematical use (if we ignore the psychological dimension of usage), and of course the grammars G and G^* are isomorphic and interchangeable. However, we have introduced the symbol $=$, distinct from $\equiv$, because we shall shortly need to consider a grammar which contains *both* of these symbols.

character, that it becomes difficult to see how students could have any difficulty in using them, formally or informally.

But hold on — do we have the right system? Does the student in high school algebra really solve his equations within a deductive system comparable to $(G^*, \vdash^*)$? If we look closely over his shoulder, we will observe that something is missing from our logical analysis of his work. Namely, while it is quite true that he is seeking to solve equations which do not hold identically, it is easy to find points of his solution where he makes use of the CAZ laws (and other identities, for example those involving the operation of negation, because he is working in a grammar possessing an operation-symbol denoting negation).

In short, the usual process of solving equations in elementary algebra involves simultaneous use of equational identities and ordinary equations. To study the underlying logic for structures of the type of $\mathfrak{N}$, we must consider a grammar G^{**} which possesses *both* the identity symbol $\equiv$ and the equals sign $=$, as well as the usual algebra of terms, $\mathfrak{T} = (T, +_T, \mathbf{0})$. Let us see how to define a semantical notion of consequence, $\models^{**}$, and a syntactical notion of derivability, $\vdash^{**}$, for the formulas of such a grammar.

When it comes to implication, there is nothing new to add to our earlier discussion, except for one small matter of notational convenience. Namely, we shall replace the notation $W^{\mathfrak{A}}(\sigma \equiv \tau)$ by the notation $W^{\mathfrak{A}}_b(\sigma \equiv \tau)$, where $b : V \to A$, with the understanding that the truth value represented by this notation is actually independent of the particular assignment-of-values-to-variables b, and is determined solely by the structure $\mathfrak{A}$ and the terms σ and τ, in just the same way as $W^{\mathfrak{A}}(\sigma \equiv \tau)$ was defined following (3.1). The purpose of this notational change is simply to be able to refer to a truth value $W^{\mathfrak{A}}_b(\phi)$ for *any* formula ϕ of the grammar G^{**}, no matter whether ϕ is an equation or an identity. With this convention we can formulate our definition of implication, $\models^{**}$, as follows.

(7.4) Definition. *Let Γ be any set of formulas of G^{**}, and ϕ any such formula. Then we have $\Gamma \models^{**} \phi$ if, and only if, $W^{\mathfrak{A}}_b(\phi) = T$ for every structure $\mathfrak{A}$ and every $\mathfrak{A}$-assignment b such that $W^{\mathfrak{A}}_b(\psi) = T$ for all $\psi \in \Gamma$.*

It is easily seen that if all formulas of Γ and ϕ are identities, then $\Gamma \models^{**} \phi$ iff $\Gamma \models \phi$, and similarly $\models^{**}$ reduces to $\models^{*}$ in case all formulas involved are ordinary equations.

Now let us turn to the question of derivability for the grammar G^{**}. We wish to give a syntactical definition of a relation $\vdash^{**}$, in terms of appropriate rules of inference, which will be sound in the sense that whenever $\Gamma \vdash^{**} \phi$ we will have $\Gamma \models^{**} \phi$, and which will be complete in the sense that $\Gamma \vdash^{**} \phi$ whenever $\Gamma \models^{**} \phi$. Obviously the first candidates to consider are the rules (R0)–(R4) and (R0)*–(R3)* which served to define sound and complete deduction relations $\vdash$ and $\vdash^*$ for the separate grammars G (identities) and G^* (equations).

Indeed, if we define a relation $\vdash'$ in the usual way from the combined set of rules (R0)–(R4) and (R0)*–(R3)* we will find easily that $(G^{**}, \vdash')$ is a sound deductive system, but it is not complete. To see the latter fact, note that we have

$$\{(v_0 +_T \mathbf{0}) \equiv v_0\} \models (v_0 +_T \mathbf{0}) = v_0, \text{ but } not\ \{(v_0 +_T \mathbf{0}) \equiv v_0\} \vdash' (v_0 +_T \mathbf{0}) = v_0. \tag{7.5}$$

The first statement of (7.5) simply expresses the obvious fact that in any structure where the distinguished element serves as an additive identity element, it can be added to any particular element of the structure without altering it. The second statement of (7.5) can be verified by observing that the rules (R0)–(R4) and (R0)*–(R3)* defining $\vdash'$ are such that if Γ is a set of *identities*, then the only *equations* ϕ such that $\Gamma \vdash' \phi$ are instances $\sigma = \sigma$ of (R0)*.

Thus it appears that if we are to achieve a derivability relation which is complete, we must strengthen the relation $\vdash'$ by adding some new rules of inference — and the example (7.5) immediately suggests the following rule.

(7.6) Rule (R5). *To infer an equation $\sigma = \tau$ from the identity $\sigma \equiv \tau$, where σ, τ are any terms.*

Now if we define a deduction relation $\vdash''$ in the usual way in terms of the rules (R0)–(R5) and

(R0)*–(R3)*, we again get a sound deductive system $(G, \vdash'')$, stronger than the system $(G, \vdash')$. But is it strong *enough*? Can we obtain a deduction to show $\Gamma \vdash'' \phi$ whenever $\Gamma \vDash \phi$? In short, is $(G, \vdash'')$ complete? Unfortunately, the answer is still no; we shall have to strengthen the relation $\vdash''$ by adding still another rule of inference. But this time we shall encounter complications.

A simple example to show the incompleteness of $\vdash''$ is provided by the following.

(7.7) $\{(\mathbf{0} +_T \mathbf{0}) = \mathbf{0}\} \vDash (\mathbf{0} +_T \mathbf{0}) \equiv \mathbf{0}$, but *not* $\{(\mathbf{0} +_T \mathbf{0}) = \mathbf{0}\} \vdash'' (\mathbf{0} +_T \mathbf{0}) \equiv \mathbf{0}$.

The first statement of (7.7) may be explained by observing that an equation and the corresponding identity have the same meaning if there are no variables occurring on either side. The second statement of (7.7) may be verified by checking that the rules (R0)–(R5) and (R0)*–(R3)* defining $\vdash''$ are such that if Γ is any set of *equations*, then the only *identities* ϕ such that $\Gamma \vdash'' \phi$ are instances $\sigma \equiv \sigma$ of (R0).

We seem to need a further rule of inference which will allow us to derive identities from equations under suitable circumstances. Of course we cannot have a general rule allowing us to pass from *any* equation $\sigma = \tau$ to the corresponding identity $\sigma \equiv \tau$; that would be much too strong, and would result in a system that was not sound. On the other hand, to limit this rule to the case where neither σ nor τ contains any variable (the case occurring in 7.7) is too weak, for in the resulting deductive system we would be unable to derive the identity $(v_1 +_T \mathbf{0}) \equiv \mathbf{0}$ from the pair of formulas

$$\Gamma_0 = \{(v_1 +_T \mathbf{0}) \equiv (v_2 +_T \mathbf{0}), \quad (v_3 +_T \mathbf{0}) = \mathbf{0}\},$$

even though we can easily show that

(7.8) $$\Gamma_0 \vDash^{**} (v_1 +_T 0) \equiv 0.$$

The solution of our difficulty turns out to be surprisingly complicated. What we must do is not simply to add a new rule of inference, but to restrict its use by adding an exceptional clause to our usual definition of the relation of derivability, as follows.

(7.9) Rule (R6). *Where σ and τ are terms, to infer the identity $\sigma \equiv \tau$ from the equation $\sigma = \tau$.*

(7.10) Definition. *Let Γ be any set of formulas of the grammar G^{**}. By a* ***derivation*** *from Γ is meant a finite sequence $(\phi_0, \ldots, \phi_n)$ of formulas of G^{**} such that each ϕ_j is either an element of Γ, or is an instance of (R0) or of (R0)*, or else is obtained from one or two preceding formulas of the sequence by one of the rules (R1)–(R6) or (R1)*–(R3)* — provided that if ϕ_j is obtained by (R6) from ϕ_k for some $k < j$, then no variable of ϕ_k shall occur in an equation of Γ which is among the formulas $\phi_0, \ldots, \phi_k$.*

We shall use the notation $\Gamma \vdash^{**} \phi$ to indicate the existence of a derivation from Γ having ϕ as its last formula; and in this case we shall say that ϕ is ***derivable*** *from Γ in the grammar G^{**}.*

Using this derivability relation $\vdash^{**}$, we can show that corresponding to (7.8) we have

(7.11) $$\Gamma_0 \vdash^{**} (v_1 +_T \mathbf{0}) \equiv \mathbf{0}$$

by means of the following 6-line derivation from Γ_0.

1.	$(v_1 +_T \mathbf{0}) \equiv (v_2 +_T \mathbf{0})$,	in Γ_0.
2.	$(v_1 +_T \mathbf{0}) \equiv (v_3 +_T \mathbf{0})$,	from line 1 by (R4).
3.	$(v_1 +_T \mathbf{0}) = (v_3 +_T \mathbf{0})$,	from line 2 by (R5).
4.	$(v_3 +_T \mathbf{0}) = \mathbf{0}$,	in Γ_0.
5.	$(v_1 +_T \mathbf{0}) = \mathbf{0}$,	from lines 3, 4 by (R2)*.
6.	$(v_1 +_T \mathbf{0}) \equiv \mathbf{0}$,	from line 5 by (R6).

Note that the sequence of formulas on lines 1–6 qualifies as a derivation from Γ_0, according to Definition (7.10), because the only variable occurring on line 5 (to which Rule (R6) is applied in passing to line 6) is v_1, and this variable does not occur in the equation on line 4 (which is in Γ).

In fact, our definition of $\vdash^{**}$ is not only adequate for the derivation of (7.11), but it is *completely* adequate, producing a deductive system $(G^{**}, \vdash^{**})$ which is both sound and complete.

(7.12) THEOREM. *If Γ is any set of formulas of the grammar G^{**} and ϕ is any such formula, then we have $\Gamma \vdash^{**} \phi$ iff $\Gamma \models^{**} \phi$.*

We are not going to prove this theorem here, which presents difficulties both in the Soundness and the Completeness parts. We content ourselves by indicating that the basic idea of the completeness proof for the grammar G, given in Section 4, must be complicated by the introduction of infinitely many new constant symbols to the grammar G^{**}, in a way suggested by one of the standard proofs of completeness for deductive systems of first-order predicate logic. (See **[2]**.) Indeed, those who have studied systems of first-order logic will recognize that our rule (R6) is essentially a form of the rule of generalization familiar from those systems, and that the restriction we have placed on using (R6) in forming deductions is familiar from systems of predicate logic.

Now, what are the pedagogical implications of our findings? It seems that the deductive system $(G^{**}, \vdash^{**})$ is quite a bit more complicated than either $(G, \vdash)$ or $(G^{*}, \vdash^{*})$, because of difficulties encountered in properly restricting the use of the rule of inference (R6). Nevertheless, it is just this more difficult system of equational logic which underlies the proof procedures normally taught in secondary schools in connection with the solution of sets of equations.[20]

Furthermore, we must keep in mind that the connection between finding solutions to equations and the logical notion of implication ($\models$), is rarely made explicit in school algebra courses — much less do we distinguish between, and compare, implication ($\models$) and derivability ($\vdash$). Finally, in this list of omissions, let us remind ourselves that the formal distinction between identities and ordinary equations, made in this paper through the use of the symbols $\equiv$ and $=$, is normally blurred by a pervasive use of the equality sign, " = ", and the student is left to guess from the context when (s)he is dealing with an identity and when (s)he has an ordinary equation.[21]

With so much that is logically relevant withheld from our students, perhaps it is easier to understand why only the bright ones sail through without a struggle. We hasten to add that we do not — definitely *not*! — advocate the study of equational logic, along the lines of this paper, for the ordinary high school algebra course. But it does seem to us that a sound understanding of this logic is a requirement for those who seek to improve our secondary curriculum.

References

1. G. Birkhoff, On the structure of abstract algebras, Proc. Cambridge Philos. Soc., 31 (1935) 433–454.

2. L. Henkin, The completeness of the first-order functional calculus, J. Symbolic Logic, 14 (1949) 159–166.

3. ———, On mathematical induction, this MONTHLY, 67 (1960) 323–338.

4. C. Martin, Equational theories of natural numbers and transfinite ordinals, Doctoral dissertation, University of California, Berkeley, 1973.

5. D. Monk, Introduction to Set Theory, McGraw-Hill, New York, 1969.

6. A. Tarski, The Concept of Truth in Formalized Languages, Logic, Semantics, Metamathematics, Clarendon Press, Oxford, 1956.

7. ———, Equational logic and equational theories of algebras, Contributions to Mathematical Logic (K. Schütte, Ed.), North-Holland, Amsterdam, 1968.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CA 94720.

[20] As the referee has pointed out, the solution of quadratic equations must rest on a still more complicated logical system, which must encompass not only equational logic, but also the logic of disjunction, "*or*". This is because we need a rule of inference allowing us to pass from the equation $\sigma \cdot_T \tau = 0$ to the disjunction $\sigma = 0$ *or* $\tau = 0$.

[21] Another way to formalize the logic of equations and identities, mentioned by the referee, is to use a single equality sign, but to use variables from two different alphabets — say lower case letters for ordinary equations and capital letters for equational identities. (One could even allow mixed equations, in which both kinds of variables occur.) Such a system, while perfectly satisfactory, seems to us somewhat further removed from ordinary school-algebra practice. In any case the rules of inference needed for a complete deductive system are no less complicated.

DEMOGRAPHIC DATING OF THE NUKUORO SOCIETY

JAMES C. FRAUENTHAL AND NOREEN GOLDMAN

Abstract. The techniques of mathematical demography are employed to determine the time of human settlement of an isolated coral atoll in the South Pacific. The ability to make this estimate is made possible by the unusual social customs practiced by the inhabitants of the atoll.

The present inhabitants of Nukuoro, a small coral atoll in the Eastern Caroline Islands of the South Pacific, are the descendants of a party of drift voyagers who landed several centuries ago. Two unusual aspects of the social customs of these people make it possible to use demographic techniques to estimate the time of arrival of the original settlers.

The first feature is that the oldest surviving member of the population was by social custom the religious ruler, i.e., the chief priest (Feeney, 1974). The second is that the only individuals whose names could be mentioned after their deaths were these religious rulers (Eilers, 1934). Consequently, although the history of the Nukuoro Society is very incomplete, the names of the past rulers have been preserved by oral tradition. The first major anthropological study of the Nukuoro Atoll (Eilers, 1934) lists the names (though not the tenure) of the 100 priests who had reigned between the time of arrival of the original settlers and the year 1871.

In a brief survey of the history of the Nukuoro, Carroll (1975) observes that prior to about 1870, virtually no contacts had occurred between the Nukuoro Society and the Western World. A few early sightings from shipboard of natives in canoes and an estimate in 1875 by Robertson (1877), who actually visited the atoll, suggest that during the nineteenth century the Nukuoro population numbered around 150 to 200 people. Archaeological evidence (Davidson, 1971) shows no indication that the population was ever any larger, and the original party of drift voyagers is thought to have numbered just 24 (Carroll, 1975). This scant evidence will be taken to mean that over the several hundred years of human habitation, the population size of the Nukuoro remained virtually constant. There is good evidence, as well as a compelling argument (Thomas, 1971), that primitive humans are capable of maintaining their population at a strictly controlled, socially optimal size.

In order to estimate the time of settlement of Nukuoro Atoll, the following questions are considered. What is the chance that a Nukuoro infant (henceforth referred to as Ego) just born will survive to become the ruler? Assuming that Ego does survive to become the ruler, what is the expected length and the variance of Ego's reign? Once the answers to these questions are known, it is straightforward to estimate the time of arrival of the original settlers on Nukuoro.

To answer the above questions only two basic pieces of demographic information are needed: the age-specific death rate for members of the population, customarily called the force of mortality, and population age distribution at the instant of Ego's birth. An obvious but important observation is that to become ruler, Ego has only to outlive all the members of the tribe alive at the instant Ego is born, since all individuals born after Ego cannot be ruler while Ego is alive. The analysis is made formally tractable by assuming that no migration occurs and that the force of mortality does not change with time and applies to all individuals in the population. (Due to the extreme isolation of the Nukuoro Atoll and the primitive life-style of the inhabitants prior to 1871, these assumptions appear to be quite reasonable.)

Begin by defining $l(x)$, which is customarily called the survivorship, to be the probability that an individual is still alive at age x. The distribution function for the random lifetime of an individual is therefore given by $1-l(x)$, and is related to the force of mortality, $\mu(x)$, by the expression

$$\mu(x) = -\frac{1}{l(x)}\frac{dl(x)}{dx}. \tag{1}$$

This equation can be integrated to give

$$l(x)=\exp\left(-\int_0^x \mu(\xi)d\xi\right), \tag{2}$$

where for consistency, the constant of integration has been chosen so that $l(0)=1$. It is convenient now to define another quantity, the expected number of years of life remaining to an individual of age t, $\mathring{e}_t$. This is readily seen to be

$$\mathring{e}_t=\frac{\int_t^\infty (x-t)\mu(x)l(x)dx}{\int_t^\infty \mu(x)l(x)dx}. \tag{3}$$

Substituting from (1), integrating the numerator by parts, and using $l(\infty)=0$ (since everyone eventually dies) leads to

$$\mathring{e}_t=\frac{1}{l(t)}\int_t^\infty l(y)dy. \tag{4}$$

Next, the probability that Ego is the ruler is determined as a function of Ego's age. Note that for Ego to be the ruler two independent events must occur. First, all K of the individuals alive at the instant of Ego's birth (the Elders) must be dead, and second, Ego must still be alive. Consider first an Elder who is age x when Ego is born at $t=0$. This Elder will then be age $x+t$ when Ego is age t, and will have probability $l(x+t)/l(x)$ of still being alive. Since the deaths of each of the Elders are independent events, the probability $P(t)$ that all the K Elders will have died by time t is

$$P(t)=\prod_{j=1}^{K}\left[1-\frac{l(x_j+t)}{l(x_j)}\right], \tag{5}$$

where the ages $x_1, x_2, \ldots, x_K$ are the ages of the Elders at the instant of Ego's birth. The product in (5) is next replaced by the sum

$$P(t)=\exp\left\{\sum_{j=1}^{K}\ln\left[1-\frac{l(x_j+t)}{l(x_j)}\right]\right\}. \tag{6}$$

Then define $K(x)$ to be the number of Elders who are younger than age x at the time of Ego's birth. The summation in (6) may then be formally replaced by an integral, thus

$$P(t)=\exp\left\{\int_0^\infty \ln\left[1-\frac{l(x+t)}{l(x)}\right]dK(x)\right\}, \tag{7}$$

where $K(\infty)=K$, the initial number of Elders. The quantity $P(t)$ is the distribution function for all of the Elders being dead by the time that Ego is age t.

The probability that Ego *is* the ruler at time t is the product of the probability that all of the Elders are dead at t and the probability that Ego is still alive; thus

$$\text{Prob}\{\text{Ego is ruler at } t\}=P(t)l(t). \tag{8}$$

But to answer the questions posed at the beginning of this paper, it is necessary to know the density, $f(t)$, for the time t at which Ego *becomes* the ruler. This is clearly equal to the product of the density for the time t at which the last surviving Elder dies and the probability that Ego is still alive at this time. But the first quantity is simply the time derivative of $P(t)$; thus

$$f(t)=\frac{dP(t)}{dt}l(t). \tag{9}$$

The probability Q that Ego ever becomes the ruler is then the integral of the expression in (9) over all possible ages t, which following one integration by parts and substitution from (1) leads to

$$Q = \int_0^\infty P(t)l(t)\mu(t)dt. \tag{10}$$

Note that the integral in (10) is the probability that Ego is ruler at the time of death. However, since no one can replace a ruler until the ruler dies, (10) is also the probability of Ego's ever becoming ruler.

The expected amount of time E that Ego survives as the ruler (given that Ego becomes the ruler) is easily found once it is recognized that Ego's expectation of life at age t from (4) is also the conditional expectation of Ego's tenure as ruler, given that coronation occurs at age t. The conditioning on the age of coronation is eliminated by multiplying (4) by the probability that Ego becomes ruler at age t from (9), and integrating over all possible ages. Following one integration by parts and normalization by Q, the result is

$$E = \frac{1}{Q}\int_0^\infty P(t)l(t)dt. \tag{11}$$

The variance V in the length of Ego's reign (given that Ego becomes the ruler) is most easily found by starting with the conditional expectation of the square of the length of Ego's tenure as ruler, given that coronation occurs at age t. This quantity follows immediately from the definition of the expectation of life at age t in (3). Proceeding as above to eliminate the conditioning, and suitably normalizing and centering about the mean lead to

$$V = \frac{2}{Q}\int_0^\infty P(t)l(t)\mathring{e}_t dt - E^2. \tag{12}$$

The standard deviation σ of the length of Ego's reign is simply the square root of V.

The time of arrival of the drift voyagers at Nukuoro atoll can be formally determined by working backwards from the year 1871 when the one hundredth ruler died. If the tenure in years of the ith ruler is denoted by T_i, then settlement occurred $T_1 + T_2 + \cdots + T_{100}$ years prior to 1871. If it is assumed that the tenures of the rulers are independently and identically distributed, with

$$E = \text{Expected}\{T_i\} \quad \text{and}$$
$$V = \text{Variance}\{T_i\} \quad \text{for all } i,$$

then the estimated time of settlement is just $100E$ years prior to 1871, and the standard deviation of this estimate is $\sqrt{100V}$ or 10σ years. It remains only to determine the values of E and V for the Nukuoro Society.

In order to estimate these quantities, a survivorship curve $l(x)$ and an age-distribution density function $k(x)$, which will approximate the age-distribution described exactly by $K(x)$ in equation (7), must be obtained. Since it has been assumed that the total population size of the Nukuoro Society has remained constant, it is consistent to assume also that the age-specific fertility and mortality have been time independent. It is then not difficult to show (for example, see Keyfitz (1968), Chapter 5) that the age distribution density function is proportional to the survivorship function. The proportionality constant is found by demanding that the total population size equals K, which leads to

$$k(x) = \frac{Kl(x)}{\int_0^\infty l(\xi)d\xi} = \frac{Kl(x)}{\mathring{e}_0}, \tag{13}$$

where $\mathring{e}_0$ is the expectation of life at age zero, and follows immediately from equation (4). Note that $K/\mathring{e}_0$ can be interpreted as being either the birth rate or the death rate for the population, which are of course the same, since the population is constant in size.

A survivorship curve must next be constructed for the Nukuoro inhabitants. Demographic and historical analyses indicate that stationarity of population size was most likely maintained with a combination of moderately low fertility and mortality (Carroll, 1975); this is in contrast to the notion that pre-industrialized European populations experienced very high mortality. Mortality data collected for the Nukuoro cohort born between 1890 and 1899 suggest that the expectation of life,

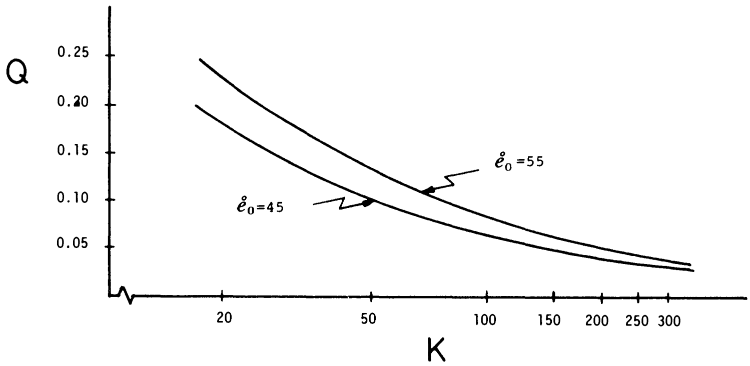


FIG. 1: Probability that Ego ever becomes the Ruler, Q, versus the total population of Elders, K. Results are shown for two different expectations of life at birth, $\mathring{e}_0$.

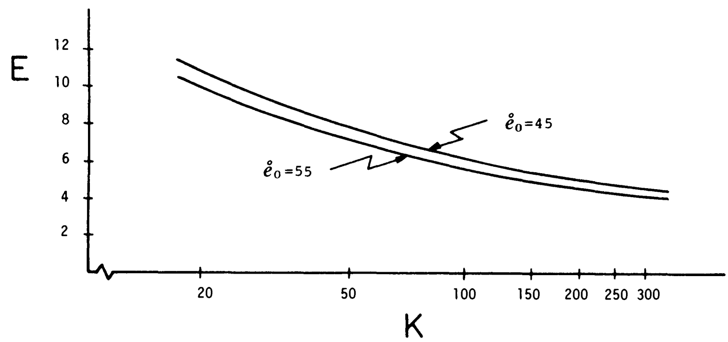


FIG. 2: Expected length of Ego's reign as Ruler in years, E, versus the total population of Elders, K. Results are shown for two different expectations of life at birth, $\mathring{e}_0$.

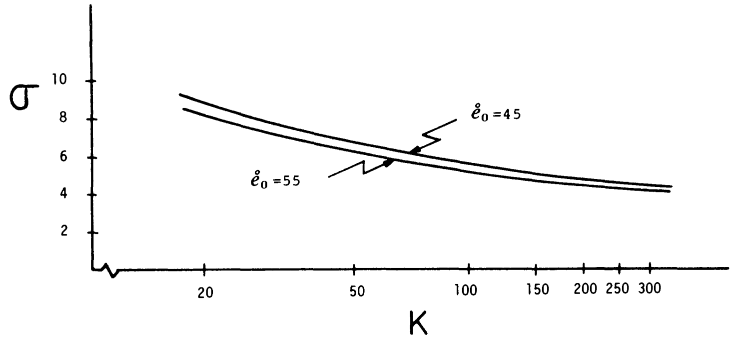


FIG. 3: Standard deviation of the length of Ego's reign as Ruler in years, σ, versus the total population of Elders, K. Results are shown for two different expectations of life at birth, $\mathring{e}_0$.

defined in equation (4), for this cohort is about 64 years. A correction for unrecorded infant mortality gives an estimate of $\mathring{e}_0$ greater than 57 years. However, this is not enough information for constructing a life table, i.e., the $l(x)$ curve described only at integer ages, for Nukuoro.

Consequently, model life tables were employed. Such tables consist of a series of life tables indexed by one or more parameters. The series is based on a wide range of recorded mortality data and each model table is in effect an average of various mortality experiences. A frequently used collection of model life tables is that derived by Coale and Demeny (1966), consisting of four "regional" sets of tables: "west," "east," "north," and "south." Each set contains a series of 24 tables indexed by $\mathring{e}_0$. Since the expectation of life is a reasonable one-parameter summary of the $\mu(x)$ or $l(x)$ function, it is frequently employed to estimate mortality curves when adequate data are not available.

Based on Nukuoro data, two model life tables ("west") were chosen for the analysis: $\mathring{e}_0 = 55$ years and $\mathring{e}_0 = 45$ years. The first table was chosen on the hypothesis that Nukuoro mortality in the last century closely approximates mortality in the last several centuries — a possibility, since the Nukuoro were an isolated people. The second table was chosen on the premise that mortality was higher in the past, but not as high as in western populations of the time.

The Coale and Demeny tables describe the $l(x)$ curve at intervals of five years of age, from age zero to age 80. In order to closely approximate the integrals in equations (10), (11) and (12), survivorship is needed by single years of age, extending into very old age. The latter condition is particularly important since mortality in the oldest ages contributes very heavily to the values of Q, E, and V.

Using a method described by Brass (1971), each model life table was fitted to a U.S. life table (1959–61) in which $l(x)$ was given by single years of age until age 109. The fit was obtained by least squares linear regression of the logits of the model $l(x)$, where

$$\mathrm{LOGIT}(l(x)) = \frac{1}{2}\ln\left(\frac{1-l(x)}{l(x)}\right),$$

on the logits of the U.S. $l(x)$, for intervals of five years of age. The fit was very satisfactory and the resulting regression line yielded a detailed model life table for Nukuoro.

Using the life table with single years of age, the integrals in equations (10), (11), and (12) were approximated by summations; the upper limit ∞ was replaced by 109, the upper age limit in the U.S. life table. $P(t)$ was obtained from equation (7) under the assumption of stationarity as in equation (13), and $\mu(x)$ derived from equation (1) by using a finite difference scheme (Kellison, 1975). Results for Q, E, and σ were derived for the two life tables described above and for values of K, the total population size (not including Ego), ranging between 20 and 300. Values for Q, E, and σ are plotted in Figures 1–3.

From the figures it is apparent that the expected length and standard deviation of Ego's tenure as ruler do not depend critically on the life table chosen. Interestingly, the expected tenure is higher for the higher mortality population: i.e., if someone in the $\mathring{e}_0 = 45$ population survives long enough to become ruler, he can expect to reign longer than a ruler in the $\mathring{e}_0 = 55$ population.

Figure 4 shows the expected date of settlement with one standard deviation confidence bounds for various values of K and $\mathring{e}_0 = 55$. As mentioned previously, these estimates are valid under the assumptions of constant population size and constant mortality throughout the period. Figures 2 and 3 illustrate that the results are largely insensitive to a 10 year range in $\mathring{e}_0$. Under the further assumption that the original party of 24 drift voyagers quickly multiplied to obtain a stationary size of 200 for the remainder of the period, an estimated year of settlement of Nukuoro is 1405, with a standard deviation of 44 years. This estimate agrees remarkably well with results from radiocarbon dating on the Nukuoro atoll. Archaeological investigations on three sites of the atoll have yielded minimum datings with one standard deviation of 1495 ± 80, 1430 ± 80, and 1530 ± 100 (Davidson, 1971).

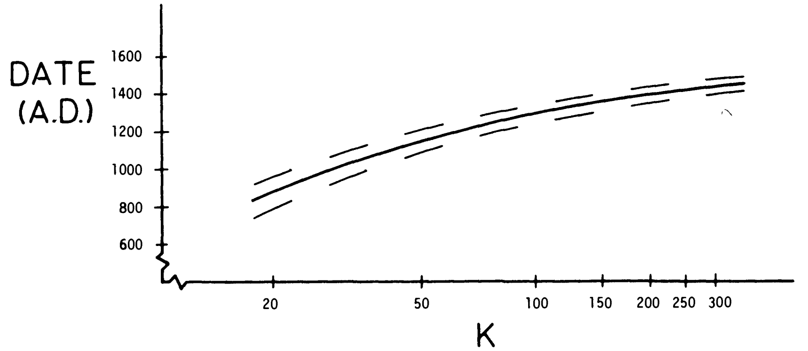


FIG. 4: Estimated date of settlement of Nukuoro Atoll versus the total population size, K. The dashed lines indicate plus and minus one standard deviation, σ. Results are shown for a population with expectation of life at birth, $\mathring{e}_0 = 55$ years.

Acknowledgements. The work presented in this paper was supported by the National Science Foundation under Grant GI-39454. The authors would like to acknowledge Griffith Feeney who suggested this problem and also solved it independently, and Nathan Keyfitz and S. Krishnamoorthy who helped us to formulate the solution.

References

1. W. Brass, On the scale of mortality, Biological Aspects of Demography, W. Brass, ed., Taylor and Francis, Ltd., London, 1971.

2. V. Carroll, The population of Nukuoro in historical perspective, Chap. 8, Pacific Atoll Populations, V. Carroll, ed., University of Hawaii, Honolulu, 1975.

3. A. Coale and P. Demeny, Regional Model Life Tables and Stable Populations, Princeton University Press, Princeton, New Jersey, 1966.

4. J. M. Davidson, Archaeology on Nukuoro Atoll, a Polynesian outlier in the Eastern Caroline Islands, Bull. No. 9, Auckland Museum and Institute, 1971.

5. A. Eilers, Inseln um Ponape. Hamburgische Wissenschaftliche Stiftung. Ergebnisse der Südsee-Expedition 1908–1910, Series 2, Subseries B., vol. 8, Friedrichsen, Degruyter, Hamburg, 1934.

6. G. Feeney, On the Oldest Person in a Stationary Population, manuscript, East-West Population Institute, University of Hawaii, Honolulu, 1974.

7. N. Keyfitz, Introduction to the Mathematics of Population, Addison–Wesley, Reading, Massachusetts, 1968.

8. S. Kellison, Fundamentals of Numerical Analysis, Richard D. Irwin, Homewood, Illinois, 1975.

9. R. Robertson, The Caroline Islands, Trans. Asiatic Soc. Japan, 5 (1877).

10. H. A. Thomas, Jr., Population dynamics of primitive societies, in Is There an Optimum Level of Population?, McGraw–Hill, New York, 1971.

APPLIED MATHEMATICS AND STATISTICS, STATE UNIVERSITY OF NEW YORK, STONY BROOK, NY 11794.
CENTER FOR POPULATION STUDIES, HARVARD UNIVERSITY, CAMBRIDGE, MA 02138.

MAGIC CUBES AND PROUHET SEQUENCES

ALLAN ADLER AND SHUO-YEN ROBERT LI

1. Introduction. Let T be a positive integer. By a magic square of order T, we mean a T by T matrix whose entries are taken from the numbers $1, 2, 3, \ldots, T^2$, and such that the sum of the entries in any row, column, or diagonal is the same number. One also requires that each of the numbers $1, \ldots, T^2$ be used exactly once as an entry. It is then a simple matter to prove that the sum of the

entries in any row, column, or diagonal of a magic square of order T is $\frac{1}{2}T(T^2+1)$. It is known that one can construct magic squares of any order except 2. (See, for example, the book *Magic Squares and Cubes*, by Andrews [**1**].)

Following Andrews, we define a magic N-cube of order T to be an N-dimensional cubical arrangement of the numbers $1, 2, 3, \ldots, T^N$ such that the sum of the numbers on any line parallel to an edge or on a *great* diagonal of the cube is the same.

In this paper, we show how to construct magic N-cubes of certain orders from generalized Morse–Hedlund sequences.

For the sake of clarity in the exposition of our results, we shall adopt a definition of magic cubes which is slightly different; however, one can see very easily that it is equivalent to our original definition.

DEFINITIONS: By an *N-cube of order T*, we mean an N-dimensional array of T^N N-dimensional boxes. A file of T boxes parallel to an edge is called an *orthogonal.* Consider a way of assigning to each box a number from the set $\{0, 1, \ldots, T^N - 1\}$ such that no two boxes have the same number. When the average of the numbers in each orthogonal or in each of the 2^{N-1} great diagonals is equal to the average of the numbers in the whole N-cube, we say that this assignment defines a *magic N-cube.*

In view of our definition of magic cubes, it is possible, and in fact convenient, to express each of its entries by an N-digit number in the base T. In that notation, the entries run from

$$\overbrace{00\ldots0}^{N} \quad \text{to} \quad \overbrace{T-1, T-1, \ldots, T-1}^{N}$$

If $T = t^M$, then the entries can also be expressed as MN-digit numbers in the base t. If, for each k satisfying $1 \leqq k \leqq MN$, the average value of the kth digit over the entries in a file is $(t-1)/2$, then the average number on this file is

$$\frac{t^{MN}-1}{2},$$

which is also the average of all of the numbers in the N-cube. Entirely similar considerations apply to the great diagonals. This is the main tool used in this paper for constructing magic squares. This method will also construct orthogonal pairs of Latin squares.

The t-ary Morse–Hedlund sequence is an infinite sequence $\{a_n\}_{n=0,1,2,\ldots}$ of integers from the set $\{0, 1, \ldots, t-1\}$. Often, it will be convenient for us to view these values as being in the integers modulo t, and to perform operations on them accordingly. One of the many equivalent definitions of the Morse–Hedlund sequence is as follows. We simply take a_n to be the sum, modulo t, of all the digits in the expression for n in base t. For instance, the procedure to obtain the ternary Morse–Hedlund sequence is described by the following table.

TABLE I

n	ternary expression	a_n
0	0	0
1	1	1
2	2	2
3	10	1
4	11	2
5	12	0
6	20	2

Continuing in this way, one obtains 012120201120201012201012. . . . . The binary Morse–Hedlund sequence is the one actually discovered by Morse and Hedlund in their works [**5**, **6**] on symbolic dynamics. It had been discovered much earlier by Thue [**9**, **10**]. However, the definition of this most general Morse–Hedlund sequence was given by E. Prouhet [**8**] in a paper in *Comptes Rendus* 1851. As this was the earliest as well as the most general treatment of these sequences, we feel that they ought to be called "Prouhet sequences," of which the Morse–Hedlund sequence is a special case. Prouhet's investigation of his sequences was apparently confined to his solution of what was later called the Tarry–Escott problem. In the last part of this paper, we give an exposition of his result as well as some generalizations which appeared in *Scripta Mathematica* in the late 1940's and early 1950's, and finally derive additional properties of the magic cubes constructed in Section 2.

2. Basic construction. In this section, we show how to construct magic N-cubes of order t^M whenever $t \mid MN$, provided that $M \geqq 1$ for t odd and $M \geqq 2$ for t even. In the course of these constructions, we shall have occasion to use the following notation.

1. t and E will denote two positive integers which will remain fixed throughout the discussion in this section.
2. If x is a non-negative integer which is less than t^E, it has an expression $x_1 \cdots x_E$ in the base t.
3. If y is another integer, we will use the expression $x \circ y$ to denote the integer whose expansion in base t is $Z_1 \cdots Z_E$ with $Z_i \equiv x_i + y \bmod t$ for $i = 1, \ldots, E$. Of course, this operation $\circ$ is not symmetric.
4. Denote by $S(t, E)$ the set $\{0, 1, \ldots, t^E - 1\}$.
5. The t-ary Morse–Hedlund sequence will still be denoted by $\{a_n\}$.

LEMMA 1. *Define the mapping $\phi: S(t, E) \to S(t, E)$ by $\phi(n) = n \circ a_n$. Then ϕ is a permutation of the set $S(t, E)$ provided that E is divisible by t.*

Proof: Let x be an element of $S(t, E)$, and let $x_1 \cdots x_E$ be its t-ary expansion. Then the t-ary expansion of $\phi(x)$ is $y_1 \cdots y_E$, where for $i = 1, \ldots, E$ we have $y_i \equiv x_i + a_x$ modulo t. Then we have $a_{\phi(x)} \equiv a_x + E \cdot a_x \equiv a_x$ modulo t, since $t \mid E$. Therefore, if we iterate ϕ, we have that $a_{\phi^\nu(x)} = a_x$ for all x and all ν, and that the k-th digit of $\phi^\nu(x)$ is congruent to $x_i + \nu a_x \bmod t$. In particular, the k-th digit of $\phi^t(x)$ is congruent to $x_i + ta_x \equiv x_i \bmod t$, so in fact $\phi^t(x) = x$ for all $x \in S(t, E)$. It follows that ϕ is a permutation. Q.E.D.

Now we consider an N-cube of order t^M with $MN = E$. *We assume throughout that t divides E and that $M \geqq 1$ when t is odd and $M \geqq 2$ when t is even.* The t^E boxes in the cube should be indexed as $0, 1, \ldots, t^E - 1$ according to the lexicographic ordering of the coordinates of the box.

THEOREM 1. *If we assign the number $\phi(n)$ to the n-th box, we obtain a magic N-cube of order t^M.*

Proof: We shall prove only the case where t is even and $M \geqq 2$. The other case is simpler and will be left to the reader.

Consider an orthogonal which is parallel to, say, the $(n+1)$st basis vector on the N-cube. The boxes on this orthogonal are indexed by an arithmetic progression

$$j + kt^{M(n+1)}, j + t^{Mn} + kt^{M(n+1)}, j + 2t^{Mn} + kt^{M(n+1)}, \ldots, j + (t^M - 1)t^{Mn} + kt^{M(n+1)},$$

where $j < t^{Mn}$.

Because $M \geqq 2$, we have $M(n+1) \geqq Mn + 2$. Therefore, adding multiples of t^{Mn} to $j + it^{Mn+2}$ will not affect any digits except for the $(Mn+1)$st and $(Mn+2)$nd. Therefore, if we break this progression up into blocks of length t^2, we obtain subsequences of the form

$$j + 0 \cdot t^{Mn} + it^{Mn+2}, j + t^{Mn} + it^{Mn+2}, j + 2 \cdot t^{Mn} + i \cdot t^{Mn+2}, \ldots, j + (t^2 - 1)t^{Mn} + it^{Mn+2}.$$

The corresponding terms of the t-ary Morse–Hedlund sequence are evenly distributed over $\{0, 1, \ldots, t-1\}$. In fact, they are

$$a, a+1, a+2, \ldots, a-1;\ a+1, a+2, \ldots, a;\ a+2, \ldots, a+1; \ldots;\ a-1, \ldots, a-2,$$

where $a = a_{j+it}Mn + 2$.

Consider the E-digit t-ary expansions of the following numbers:

$$\phi(j + 0 \cdot t^{Mn} + i \cdot t^{Mn+2}), \ldots, \phi(j + (t^2 - 1)t^{Mn} + i \cdot t^{Mn+2}).$$

We want to show that their kth digits are evenly distributed modulo t for every k, since that will immediately imply that the average of the entries in an orthogonal is $(t^{MN} - 1)/2$. Except for $k = (MN + 1)$ and $k = (MN + 2)$, this is completely obvious. For $k = (MN + 1)$, the sequence of kth digits in a block of length t^2 is given by

$$a, a+2, a+4, \ldots, a-2;\ a+1, a+3, a+5, \ldots, a-1;\ a+2, \ldots, a; \ldots; a-1,\ a+1, \ldots, a-3.$$

For $k = (MN + 2)$, the sequence of kth digits in a block of length t^2 is given by

$$a, a+1, a+2, \ldots, a-1;\ a+2, a+3, \ldots, a+1;\ a+4, \ldots, a+3; \ldots; a-2,\ a-1, a, \ldots a-3.$$

These are easily seen to be evenly distributed. In view of our remarks following the definition of magic N-cubes, this proves that the average of the entries on an orthogonal is $\frac{1}{2}(t^E - 1)$.

We shall now prove that the average of the entries on a great diagonal is also $\frac{1}{2}(t^E - 1)$, which will complete the proof that our cube is magic. In order to facilitate our discussion, let us call two E-digit t-ary numbers *complementary* in case they add up to $t^E - 1$. This will be so if and only if for every k, their kth digits add up to $t - 1$.

If t is even, one can take the boxes on a great diagonal in pairs whose indices are complementary. When t is odd, one can take them in such pairs except for the box in the middle which is left over. If (x, y) is a pair of complementary indices of boxes on a given great diagonal, we have $a_x + a_y \equiv E(t-1) \equiv 0 \pmod t$ since t divides E. Therefore, $a_x \equiv -a_y \bmod t$, so that $\phi(x)$ and $\phi(y)$ will also be complementary. If t is odd, $a_x = 0$ for $x = \frac{1}{2}(t^E - 1)$, so that $\phi(\frac{1}{2}(t^E - 1)) = \frac{1}{2}(t^E - 1)$. This completes the proof. QED

In the following examples, we have added 1 to all of the entries of our magic N-cubes so that they will run from 1 up to t^{MN} as is customary.

Examples:

(1) The $3 \times 3 \times 3$ magic cube constructed by this method is (after adding one to each entry)

1	15	26
17	19	6
24	8	10

23	7	12
3	14	25
16	21	5

18	20	4
22	9	11
2	13	27

FIG. 1.

(2) The (adjusted) magic square of order 5 which is constructed by this method is:

1	8	15	17	24
12	19	21	3	10
23	5	7	14	16
9	11	18	25	2
20	22	4	6	13

FIG. 2

(3) In case $t = 2$, the procedure for making magic N-cubes of order 2^M can be considerably simplified and could, at least for $N = 2, 3$ easily be described to a junior high school class.

RECIPE:

(i) Generate the Morse–Hedlund sequence to as many terms as there are boxes in the square or cube one wishes to make. For example, for a 4×4 magic square, we would write

$$0110100110010110.$$

(ii) Count out the terms of the sequence and write down below each 1 the number of the place where it occurs. Thus, for the example we are considering, we have

$$\begin{array}{cccccccccccccccc} 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ \hline & 2 & 3 & & 5 & & & 8 & 9 & & & 12 & & 14 & 15 & \end{array}$$

(iii) Now count through the boxes in the square, and whenever the number you count is one of the numbers which you have written below the Morse–Hedlund sequence, write the number down in the box. Otherwise, skip it and go to the next box. Thus, we get

	2	3	
5			8
9			12
	14	15	

To get the rest of the numbers in the square, count down in from 16 to 1 and fill in all the boxes that have been left out. We get

16	2	3	13
5	11	10	8
9	7	6	12
4	14	15	1

It is perhaps worth remarking that if one reflects the entries in the square through the vertical axis, one obtains the magic square which appears in Albrecht Dürer's woodcut Melancholia.

The magic N-cubes constructed by our method have a number of unexpected properties. One of these is given by the following theorem.

THEOREM 2. *In the square constructed in Theorem* 1, *the sum of the squares of the entries of an orthogonal is equal to the sum of the squares of the entries on the complementary orthogonal.*

Proof: By the complementary orthogonal, we mean the orthogonal obtained by reflecting the given orthogonal through the center of the N-cube. It is the same thing to say that the indices of its boxes are complementary to the indices on the original orthogonal. As we observed in the proof of Theorem 1, if x and y are complementary, so are $\phi(x)$ and $\phi(y)$. Accordingly, if we denote by u the number $t^E - 1$, and by $\phi_1, \ldots, \phi_{t^n}$ the entries of some orthogonal, then the entries of the complementary orthogonal are given by $u - \phi_1, \ldots, u - \phi_{t^M}$. Then

$$\left(\sum_{j=1}^{t^N} \phi_j^2\right) - \left(\sum_{j=1}^{t^N} (u - \phi_j)^2\right) = u \sum_{j=1}^{t^N} (2\phi_j - u) = 0,$$

since the average of the entries on the orthogonal is $\frac{1}{2}u$. That proves Theorem 2. Q.E.D.

When we discuss the Tarry–Escott problem in section 4, we shall derive additional properties of our magic cubes which considerably extend the results of Theorem 2.

Example (2) above shows that our construction is valid under more general conditions than those assumed in Theorem 1. Note, however, that Theorem 2 fails for this example.

3. Operations on magic N-cubes. In this section we describe some relations between magic cubes of different sizes and dimensions. Many of our results are empirical observations which have not yet been given a general formulation.

If we use the first sixty-four terms of the binary Morse–Hedlund sequence to construct a $4\times4\times4$ magic cube, we obtain

1	63	62	4
60	6	7	57
56	10	11	53
13	51	50	16

48	18	19	45
21	43	42	24
25	39	38	28
36	30	31	33

32	34	35	29
37	27	26	40
41	23	22	44
20	46	47	17

49	15	14	52
12	54	55	9
8	58	59	5
61	3	2	64

FIG. 3

We can also use the first sixty-four terms to construct an 8×8 magic square, as in Fig. 4. One sees immediately that it is obtained from the magic cube of Fig. 3 by taking the rows of the cube one at a time and using them to fill up the 8×8 square. If one uses the columns of each square of the cube instead, one obtains a completely different magic square (Fig. 5), or if one decomposes the cube in the remaining direction, one obtains yet another magic square (see Fig. 6).

1	63	62	4	60	6	7	57
56	10	11	53	13	51	50	16
48	18	19	45	21	43	42	24
25	39	38	28	36	30	31	33
32	34	35	29	37	27	26	40
41	23	22	44	20	46	47	17
49	15	14	52	12	54	55	9
8	58	59	5	61	3	2	64

FIG. 4

1	62	48	19	32	35	49	14
60	7	21	42	37	26	12	55
56	11	25	38	41	22	8	59
13	30	36	31	20	47	61	2
63	4	18	45	34	29	15	52
6	57	43	24	27	40	54	9
10	53	39	28	23	44	58	5
51	16	30	33	46	17	3	64

FIG. 5

1	48	32	49	63	18	34	15
62	19	35	14	4	45	29	52
60	21	37	12	6	48	27	54
7	42	26	55	57	24	40	9
56	25	41	8	10	39	23	58
11	38	22	59	53	28	44	5
13	36	20	61	51	30	46	3
50	31	47	2	16	33	17	64

FIG. 6

In fact, no matter how one decomposes the cube into squares, if one then uses the rows of columns of these squares to fill out an 8×8 square, one obtains a magic square. On the other hand, if one reverses this procedure with one of the squares so obtained and tries to take each half-row of one of these squares as the row of a $4\times4\times4$ cube, then one obtains new magic cubes.

For example, if one applies this to (Fig. 6) one obtains the magic cube of Fig. 7.

1	48	32	49
63	18	34	15
62	19	35	14
4	45	29	52

60	21	37	12
6	43	27	54
7	42	26	55
57	24	40	9

56	25	41	8
10	39	23	58
11	38	22	59
53	28	44	5

13	36	20	61
51	30	46	3
50	31	47	2
16	33	17	64

FIG. 7

We now indicate another operation on magic N-cubes. Suppose $A = (a_{ij})$ is an $m \times m$ magic square and $B = (b_{kl})$ is an $n \times n$ magic square. Then we can form a new magic square $C = (c_{rs})$ by the rule $c_{rs} = a_{ij} + m^2(b_{kl} - 1)$ for

$$r = i + (k-1)m, \qquad s = j + (l-1)n.$$

Then we write $C = A \times B$. If we denote by S the set of all magic squares, then $\times$ is an operation on S. It is easy to see that it is not commutative.

THEOREM 3. *$\times$ is an associative operation on S.*

Proof: Let A, B, C and S be square matrices, and say A is $m \times m$, B is $n \times n$, and C is $p \times p$, $A = (a_{ij})$, $B = (b_{kl})$, $C = (c_{\alpha\beta})$. Let $E = (A \times B) \times C$ and $F = A \times (B \times C)$, and say $E = (e_{rs})$, $F = (f_{\rho\sigma})$. Then

$$e_{rs} = (a_{ij} + m^2(b_{kl} - 1)) + (mn)^2(c_{\alpha\beta} - 1)$$

for $r = (i + (k-1)m) + (\alpha - 1)mn$, $s = (j + (l-1)m) + (\beta - 1)mn$ and

$$f_{\rho\sigma} = a_{ij} + m^2((b_{kl} + n^2(c_{\alpha\beta} - 1)) - 1)$$

for $\rho = i + ((k + (\alpha-1)n) - 1)m$, $\sigma = j + ((l + (\beta - 1)n) - 1)m$. Hence, for given $i, j, k, l, \alpha, \beta$, we have $\rho = r$, $\sigma = s$ and $e_{rs} = f_{\rho\sigma}$. Q.E.D.

Thus, (S, x) is a semigroup. The 1×1 magic square $\boxed{1}$ is a two-sided unit, and it is easy to see that the group of symmetries of the square acts as a group of automorphisms of S. It is also not difficult to show that S is cancellative, i.e., if $A \times B = A \times C$, then $B = C$, and if $B \times A = C \times A$, then $B = C$. We suspect that (S, x) is actually a free semigroup, but we have not yet proved this. These observations and conjectures extend in an obvious way to N-cubes.

If A, B are the magic squares given by

$A =$

8	1	6
3	5	7
4	9	2

and $B =$

1	15	14	4
12	6	7	9
8	10	11	5
13	3	2	16

then $A \times B$ is given by Fig. 8.

4. The Tarry-Escott Problem. The problem of finding disjoint sets $\{a_1, \ldots, a_r\}$ and $\{b_1, \ldots, b_r\}$ of distinct integers such that

$$\begin{gathered}
a_1 + a_2 + \cdots + a_r = b_1 + b_2 + \cdots + b_r,\\
a_1^2 + a_2^2 + \cdots + a_r^2 = b_1^2 + b_2^2 + \cdots + b_r^2,\\
a_1^3 + \cdots + a_r^3 = b_1^3 + \cdots + b_r^3,\\
\vdots\\
a_1^n + \cdots + a_r^n = b_1^n + \cdots + b_r^n
\end{gathered}$$

is an old one. Goldbach and Euler each made contributions to the problem. In Eugene Dickson's *Theory of Numbers Vol. II*, Chapter 24 [2], one can find a long account of various results people had

8	1	6	134	127	132	125	118	123	35	28	33
3	5	7	129	131	133	120	122	124	30	32	34
4	9	2	130	135	128	121	126	119	31	36	29
107	100	105	53	46	51	62	55	60	80	73	78
102	104	106	48	50	52	57	59	61	75	77	79
103	108	101	49	54	47	58	63	56	76	81	74
71	64	69	89	82	87	98	91	96	44	37	42
66	68	70	84	86	88	93	95	97	39	41	43
67	72	65	85	90	83	94	99	92	40	45	38
116	109	114	26	19	24	17	10	15	143	136	141
111	113	115	21	23	25	12	14	16	138	140	142
112	117	110	22	27	20	13	18	11	139	144	137

FIG. 8

obtained. Most of these results consist of special numerical solutions for $n = 2$ and 3 and parametric solutions for small values of n. The problem, of course, is to try to make n as large as possible compared with r. The best results along these lines all give values of n which are of the order of log r. On the other hand, n must be smaller than r. For if $n \geqq r$, then the sets $\{a_1, \ldots, a_r\}$ and $\{b_1, \ldots, b_r\}$ must be equal. It would be desirable to determine the asymptotic behavior of the bound of n as a function in r.

According to Dickson, the first general result giving n the order of log r is due to Prouhet [3] in a note which appeared in *Comptes Rendus de l'Académie des Sciences* in 1851. He stated that one can, for every t and n, divide the integers $1, 2, 3, \ldots, t^{n+1}$ into t disjoint sets with t^n elements each such that any two of these sets have the same power sums for any exponent less than or equal to n. He gave the following construction, although he did not prove that it had the properties which he stated. Let $\{a_m\}$ denote the t-ary Morse–Hedlund sequence. For $j = 0, \ldots, t-1$, let A_j denote the set of all $m < t^{n+1}$ for which $a_m = j$. Then the sets $A_0, A_1, \ldots, A_{t-1}$ have the properties stated by Prouhet. The first proof of this result appeared as a special case in the following theorem given by Lehmer [3].

THEOREM 4. *Let t, E, k, $\mu_1, \ldots, \mu_E$ be positive integers such that $t > 1$, $1 \leqq k < E$. Write*

$$\sigma_k(y) = \sum_{\substack{b_1 + \cdots + b_E \equiv y \pmod t \\ 0 \leqq b_1, \ldots, b_E < t}} (b_1\mu_1 + \cdots + b_E\mu_E)^k.$$

Then $\sigma_k(0) = \sigma_k(1) = \cdots = \sigma_k(t-1)$.

In particular when $\mu_k = t^{k-1}$ for $0 \leqq k \leqq E$, Theorem 4 reduces to Prouhet's result. Alternative proofs of Theorem 4 are contained in [4] and [7].

COROLLARY. *Let c be a fixed number. In the definition of $\sigma_k(y)$ in Theorem 4, if $b_1\mu_1 + \cdots + b_E\mu_E$ is replaced by $b_1\mu_1 + \cdots + b_E\mu_E + c$, then the conclusion of the theorem still holds.*

From this, we shall derive further magic properties of the magic cubes constructed in Theorem 1.

Take an N-cube of order t^M with the boxes in it indexed as $0, 1, \ldots, t^{MN} - 1$ according to the lexicographic ordering of the coordinates of the box. Consider an n-dimensional subcube S in this N-cube. Assume, say, that this subcube is parallel to the m_1th, m_2th, $\ldots, m_n$th basis vectors of the

N-cube. Let

$$J = \bigcup_{j=1}^{n} \{Mm_j, Mm_j - 1, \dots, Mm_j - t + 1\}.$$

Then the indices of the boxes in S are of the form

$$\sum_{i \in J} x_i t^{i-1} + \sum_{\substack{1 \leq i \leq MN \\ i \notin J}} c_i t^{i-1},$$

where $x_i \in \{0, 1, \dots, t-1\}$. Here the c's are constants and they uniquely determine the subcube S. We observe the following two facts.

(I) Identify every box with its index so that S is considered as a set of t^{Mn} integers. Write

$$S_y = \{s \in S \mid a_s \equiv y \mod t\}.$$

From the corollary of Theorem 4, we know

$$\sum_{s \in S_0} s^k = \sum_{s \in S_1} s^k = \cdots = \sum_{s \in S_{t-1}} s^k,$$

for $k = 0, 1, \dots, Mn - 1$.

(II) Recall the mapping ϕ defined on $\{0, 1, \dots, t^{MN} - 1\}$ via $\phi(x) = x \circ a_x$. If $s \in S_y$, then $\phi(s)$ is equal to $s \circ y$ and therefore is the index of a box on an n-dimensional subcube parallel to the n-cube S which corresponds to $c_i + y \mod t$, where $1 \leq i < MN$ and $i \notin J$. We shall call this parallel subcube a *conjugate* of S in the original N-cube. Denote this conjugate by $S \circ y$. It then follows from the above observation that

$$\sum_{s \in S} s^k = \sum_{s \in S \circ 1} s^k = \cdots = \sum_{s \in S \circ (t-1)} s^k,$$

for $k = 0, 1, \dots, Mn - 1$. We thus arrive at the following conclusion.

THEOREM 5. *In the magic N-cube of order t^M given by Theorem 1, conjugate n-dimensional subcubes have equal sum of the k-th power of their entries, $k = 0, 1, \dots, Mn - 1$.*

For instance, the magic N-cube of order t can be sliced into t pieces consisting of conjugate $(N-1)$-dimensional subcubes; and these subcubes have equal sums of like powers of their entries.

Theorem 5 is of particular interest in the special case that $t = 2$. Because then two subcubes are conjugate to each other only when they are *complementary* to each other. A one-dimensional subcube means simply an orthogonal. So in this case we may combine Theorem 2 with Theorem 5 and find that complementary orthogonals have equal sums of the kth powers of their entries for all $k \leq \max\{2, M-1\}$. This fact is generalized by the following:

THEOREM 6. *In the magic N-cube of order 2^M constructed in Theorem 1, complementary n-dimensional subcubes have equal sums of the (Mn)th powers of their entries, provided that Mn is an even integer.*

This theorem is a direct consequence of Theorem 5 and the following elementary

LEMMA. *Let c be a constant. Consider the polynomials $f_k(z) = (c-z)^k - z^k$. Then f_{2m} is a linear sum of $f_1, f_2, \dots, f_{2m-1}$.*

Proof. $f_{2m} - \binom{m}{1} c f_{2m-1} + \binom{m}{2} c^2 f_{2m-2} - \cdots + \binom{m}{m} (-c)^m f_m$

$$= (c-z)^{2m} - \binom{m}{1} c(c-z)^{2m-1} + \binom{m}{2} c^2(c-z)^{2m-2} - \cdots + \binom{m}{m}(-c)^m (c-z)^m$$
$$- \left[z^{2m} - \binom{m}{1} c z^{2m-1} + \binom{m}{2} c^2 z^{2m-2} - \cdots + \binom{m}{m}(-c)^m z^m \right]$$
$$= (c-z-c)^m (c-z)^m - (z-c)^m z^m$$
$$= 0 \qquad \text{Q.E.D.}$$

Using the lemma, we can conclude the proof of Theorem 6 as follows. Let S and $\bar{S}$ be complementary n-dimensional subcubes. Let $c = t^{Mn} - 1$. Using the notations in the lemma,

$$\sum_{z \in S} z^{Mn} - \sum_{w \in \bar{S}} w^{Mn} = \sum_{z \in S} (z^{Mn} - (t^{Mn} - 1 - z)^{Mn}) = \sum_{z \in S} f_{Mn}(z).$$

Since Mn is even, we can write $f_{Mn}(z)$ as a linear combination $\sum_{i=1}^{Mn-1} b_i f_i(z)$ so that

$$\sum_{z \in S} f_{Mn}(z) = \sum_{z \in S} \sum_{i=1}^{Mn-1} b_i f_i(z) = \sum_{i=1}^{Mn-1} b_i \sum_{z \in S} f_i(z) = 0$$

by Theorem 5.

References

1. W. S. Andrews, Magic Squares and Cubes, Dover, New York 1960.
2. L. E. Dickson, History of the Theory of Numbers, Chelsea, New York vol 2, 1952, p. 705.
3. D. H. Lehmer, The Tarry–Escott Problem, Scripta Math., 13 (1947) 37–41.
4. L. Mirsky, A remark on D. H. Lehmer's solution of the Tarry–Escott Problem, Scripta Math., 14 (1948) 126–127.
5. M. Morse, Recurrent geodesics on a surface of negative curvature, Trans. Amer. Math. Soc., 22 (1921) 84–100.
6. M. Morse and G. A. Hedlund, Unending chess, symbolic dynamics and a problem in semigroups, Duke Math. J., 11 (1944) 1–7.
7. G. Palamà, On a theorem of D. H. Lehmer concerning the Tarry–Escott Problem, Scripta Math., 19 (1953) 19–23.
8. E. Prouhet, Mémoire sur quelques relations entre les puissances des nombres, Inst. France, Acad. d. Sci., Comptes Rendus, 33 (1851) 225.
9. A. Thue, Über unendliche Zeichenreihen, Norske Vid. Selsk. Skr., I Mat.-Nat. Kl., Christiania, No. 7 (1906) 1–22.
10. ———, Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, *ibid.*, No. 1 (1912) 1–67.

School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540.
Department of Mathematics, University of Illinois, Chicago Circle, Chicago, IL 60680.

CORRECTIONS AND ADDITIONS TO "SOME APPLICATIONS OF THE WREATH PRODUCT CONSTRUCTION"

(This Monthly, 83(1976) 317–338)

Charles Wells

1. There is one important systematic error. In the proof of Theorem 13.1, every occurrence (I count 7 of them) of the phrase "$s \in I^1$" should read "$s \in I$" (delete the superscript 1). By the way, I should have stated that the presentation of this proof closely follows Lallement **[71]**.

2. The proof of the Krohn-Rhodes Theorem in §14 is (as far as I know) correct, but the inductive step is incorrectly described on p. 331. What is shown is that S is decomposable into a wreath product of semigroup actions, each of which has smaller measure *or else is a group action dividing S.*

3. Two books and an article relevant to the subject were published more or less simultaneously with the article. **[1]** is an exposition of the theory of characters of wreath products with applications to representations theory and to combinatorics. **[2]** constitutes a thorough reworking of the algebraic side of automaton theory and contains among other things two proofs of the Krohn-Rhodes Theorem, one a simpler version of the proof given in §14 of my article, and the other a reformulation of the Zeiger-Ginsburg proof which I mentioned on p. 330. This reformulation yields a stronger theorem [**2**, Theorem 7.1] which makes possible an efficient decomposition algorithm. **[3]** gives a new algebraic construction of the real numbers directly from the integers.

References

1. A. Kerber, Representations of Permutation Groups II, Lecture Notes in Math. 495, Springer-Verlag, New York, 1975.

2. S. Eilenberg, Automata, Languages and Machines, Vol. B, Academic Press, New York, 1976.

3. F. Faltin, N. Metropolis, B. Ross, and G.-C. Rota, The real numbers as a wreath product, Advances in Math., 16 (1975) 278–304.

DEPARTMENT OF MATHEMATICS AND STATISTICS, CASE WESTERN RESERVE UNIVERSITY, CLEVELAND, OH 44106.

THE TEN COMMANDMENTS OF STATISTICAL INFERENCE

MICHAEL F. DRISCOLL

The original version of these commandments has apparently been lost, perhaps in antiquity. There may now exist several variants. One has appeared in Thomas **[1]**; here is another.

I. Thou shalt not hunt statistical significance with a shotgun.
II. Thou shalt not enter the valley of the methods of inference without an experimental design.
III. Thou shalt not make statistical inference in the absence of a model.
IV. Thou shalt honor the assumptions of thy model.
V. Thou shalt not adulterate thy model to obtain significant results.
VI. Thou shalt not covet thy colleague's data.
VII. Thou shalt not bear false witness against thy control-group.
VIII. Thou shalt not worship the 0.05 significance level.
IX. Thou shalt not apply large-sample approximations in vain.
X. Thou shalt not infer causal relationship from statistical significance.

Reference

1. D. H. Thomas, Figuring Anthropology: First Principles of Probability and Statistics, Holt, Rinehart, and Winston, New York, 1976, pp. 458–468.

DEPARTMENT OF MATHEMATICS, ARIZONA STATE UNIVERSITY, TEMPE, AZ 85281.

MATHEMATICAL NOTES

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Department of Mathematics, University of Wisconsin, Madison, WI 53706.

INFINITE PRODUCTS FOR k-th ROOTS

N. J. FINE

The following elegant expansion is due to Cantor [1]:

$$\sqrt{\frac{x+1}{x-1}} = \left(1+\frac{1}{q_1}\right)\left(1+\frac{1}{q_2}\right)\cdots, \tag{1}$$

where $q_1 = x > 1$ and $q_{n+1} = 2q_n^2 - 1$ for $n \geqq 1$. There are similar expansions and generalizations by Engel [2], Ostrowski [3], Lotockiĭ [4], and Sierpiński [5]. If δ_n is the error committed in (1) by using the first n factors only, then $\delta_{n+1} < c\delta_n^2$; that is, the convergence is *quadratic.* (Many computational algorithms are of this type.) For $x = 2$, the use of 12 factors in (1) would yield $\sqrt{3}$ correct to 1,000 decimals. It is the purpose of this note to give a similar algorithm for kth roots in which the convergence is *cubic*: $\delta_{n+1} < c\delta_n^3$. The result for $k = 2$ (Eq. 9) is particularly simple.

Let z be any number greater than 1, and let a and b be positive numbers such that $a + b = k$, an integer greater than 1. Then we can write $z = (x + a)/(x - b)$, with $x > b$, and

$$\sqrt[k]{z} = \sqrt[k]{\frac{x+a}{x-b}} \approx 1 + \frac{1}{x},$$

the approximation being good for x large. Our plan is to write

$$\sqrt[k]{\frac{x+a}{x-b}} = \left(1+\frac{1}{x}\right)\sqrt[k]{\frac{y+a}{y-b}}. \tag{2}$$

If x is large, y will be even larger, and we may hope that an iteration of this procedure will lead to a rapidly convergent product. Thus, with $x_1 = x$,

$$\sqrt[k]{\frac{x+a}{x-b}} = \left(1+\frac{1}{x_1}\right)\left(1+\frac{1}{x_2}\right)\cdots\left(1+\frac{1}{x_n}\right)\sqrt[k]{\frac{x_{n+1}+a}{x_{n+1}-b}}, \tag{3}$$

where x_{n+1} is the same function of x_n as y is of x in (2). Solving for y in (2), we get

$$y = \frac{bx^k(x+a) + a(x+1)^k(x-b)}{x^k(x+a) - (x+1)^k(x-b)}, \tag{4}$$

$$y - x = \frac{(x+a)(x-b)[(x+1)^k - x^k]}{x^k(x+a) - (x+1)^k(x-b)}. \tag{5}$$

The sequence $\{x_n\}$ will be monotonic increasing if the denominator in (5) is positive for all $x > b$. It is not difficult to show that this is the case if $b \geqq (k-1)/2$. If $\{x_n\}$ is bounded it will converge to a value ξ which makes the right side of (5) vanish if x is replaced by ξ. Since this is impossible, $x_n \to \infty$ and the

corrective factor in (3) converges to 1. Hence, with $x_1 = x > b \geqq (k-1)/2$ and

$$x_{n+1} = \frac{bx_n^k(x_n+a) + a(x_n+1)^k(x_n-b)}{x_n^k(x_n+a) - (x_n+1)^k(x_n-b)} \qquad (n \geqq 1), \tag{6}$$

we have the expansion

$$\sqrt[k]{\frac{x+a}{x-b}} = \left(1+\frac{1}{x_1}\right)\left(1+\frac{1}{x_2}\right)\cdots. \tag{7}$$

Now the relative error ε_n after n factors is easy to estimate:

$$1+\frac{1}{x_{n+1}} < 1+\varepsilon_n = \sqrt[k]{\frac{x_{n+1}+a}{x_{n+1}-b}} < 1+\frac{1}{x_{n+1}-b}. \tag{8}$$

Thus $\varepsilon_n \approx x_{n+1}^{-1}$, and we should try to make x_n increase as rapidly as possible. The numerator in (6) is of degree $k+1$ in x_n, whereas the denominator is of degree $k-1$ at most, the coefficient of x_n^{k-1} being $k(b-(k-1)/2)$. Therefore quadratic convergence is assured in any case. But if we choose $b=(k-1)/2$, $a=(k+1)/2$, we get cubic convergence, with

$$x_{n+1} \approx \frac{12}{k^2-1} x_n^3.$$

For the case $k=2$, a slight change in notation yields

$$\sqrt{\frac{x+3}{x-1}} = \left(1+\frac{2}{q_1}\right)\left(1+\frac{2}{q_2}\right)\cdots, \tag{9}$$

where $q_1 = x > 1$, $q_{n+1} = q_n^3 + 3q_n^2 - 3$. Here

$$\frac{2}{q_{n+1}} < \varepsilon_n < \frac{2}{q_{n+1}-1}.$$

If $q_n \geqq 2$, say, we have $\varepsilon_n < 2q_n^{-3} \leqq 2(1/x)^{3^n}$. For $x=3$, we get

$$\sqrt{3} = \left(1+\frac{2}{3}\right)\left(1+\frac{2}{51}\right)\left(1+\frac{2}{140{,}541}\right)\cdots. \tag{10}$$

The first three factors alone would give an accuracy of 14 decimals, and twelve factors would give well over 300,000 correct decimals.

References

1. G. Cantor, Gesammelte Abhandlungen, p. 43.

2. F. Engel, Verhandl. d. 52 Versammlung deutscher Philologen and Schulmänner in Marburg, 1913.

3. A. Ostrowski, Über einige Verallgemeinerungen des Eulerschen Produktes..., Verhandl. der naturforsch. Gesellsch. in Basel, Band 40 (1929).

4. A. V. Lotockiĭ, On a method of representing real numbers as infinite products, Ivanov. Gos. Ped. Inst. Uč. Zap. Fiz.-Mat. Fak. no. 1, 1 (1941) 27–35.

5. W. Sierpiński, Généralisation d'une formule de E. B. Escott pour les racines carrées, Bull. Soc. Roy. Sci. Liège, 22 (1953) 520–529.

DEPARTMENT OF MATHEMATICS, PENNSYLVANIA STATE UNIVERSITY, UNIVERSITY PARK, PA 16802.

DOES EVERY RIGHT IDEAL OF A MATRIX RING CONTAIN A NILPOTENT ELEMENT?

G. F. BIRKENMEIER AND R. P. TUCCI

Introduction. Throughout this note R will denote an associative ring with unity, R_n will be the $n \times n$ matrix ring over R, and E_{ij} will denote the matrix with unity in the ith row and jth column and zeros elsewhere. We say R is *densely nil* (DN) if every nonzero right ideal contains a nonzero nilpotent element. Equivalently, R is DN if for every nonzero $x \in R$ there exists $r \in R$ such that $xr \neq 0$, but $(xr)^2 = 0$.

It can be shown that every nonzero two-sided ideal of R_n contains a nonzero nilpotent element. The question naturally arises as to whether we can generalize this statement to right ideals. We show that for a ring R every nonzero idempotent generated right ideal of R_n contains a nonzero nilpotent element. Furthermore, there is a large class of matrix rings which are DN. For example, matrix rings over the following types of rings are DN: right self-injective, von Neumann regular, commutative, DN, or left Ore domains. Surprisingly, matrix rings over domains which are not left Ore domains are not DN.

PROPOSITION 1. *Let R be generated (as a right ideal) by nilpotent elements. Then every nonzero idempotent generated right ideal of R contains a nonzero nilpotent element.*

Proof. Assume $e = e^2 \in R$ such that eR contains no nonzero nilpotent elements. Let $x \in R$. Then $(exe - ex)^2 = 0$, so $ex = exe$, and, by induction, $(ex)^n = ex^ne$. If, in addition, x is nilpotent, we obtain $ex = 0$, so $x \in (1-e)R$. Thus $(1-e)R = R$, and $eR = 0$.

COROLLARY 2. *Let R be a ring which is not necessarily generated by nilpotent elements. Then every nonzero idempotent generated right ideal of R_n $(n > 1)$ contains a nonzero nilpotent element.*

Proof. R_n is generated by its nilpotent elements.

The condition that every nonzero idempotent generated right ideal contains a nonzero nilpotent element is equivalent to the DN condition for many types of rings. In particular, these conditions are equivalent for von Neumann regular rings, right self-injective rings, or rings with essential socle [**2**]. Thus if R is right self-injective or von Neumann regular, then R_n $(n > 1)$ is DN because R_n is right self-injective or von Neumann regular, respectively [**1**; pp. 262 and 265].

LEMMA 3. *If R_n is DN, then R_{n+1} is DN $(n \geqq 1)$.*

Proof. Let M be a nonzero matrix of R_{n+1}. By multiplying M by the appropriate $E_{i,n+1}$, we see that MR_{n+1} contains a nonzero matrix A of the form

$$\begin{vmatrix} 0 & \cdots & 0 & a_1 \\ 0 & & 0 & a_2 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & a_{n+1} \end{vmatrix}$$

If $a_{n+1} = 0$, then A is nilpotent. If $a_{n+1} \neq 0$, then since R_n is DN there exists $Y \in R_n$ such that

$$\begin{vmatrix} 0 & \cdots & 0 & a_2 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & a_{n+1} \end{vmatrix} Y = X \neq 0 \quad \text{but} \quad X^2 = 0. \quad \text{Hence,} \quad A \begin{vmatrix} 0 & \cdots & 0 \\ \vdots & Y & \\ 0 & & \end{vmatrix}$$

is a nonzero nilpotent element of MR_{n+1}.

LEMMA 4. *R_2 is DN if and only if for every a, $b \in R$, where a and b are not both zero, there exist x, $y \in R$ such that*

$$\begin{vmatrix} ax & ay \\ bx & by \end{vmatrix} \neq \begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix}, \quad \begin{vmatrix} axax + aybx & axay + ayby \\ bxax + bybx & bxay + byby \end{vmatrix} = \begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix}. \tag{1}$$

Proof. Assume R_2 is DN. Let

$$A = \begin{vmatrix} a & 0 \\ b & 0 \end{vmatrix}.$$

Then the right ideal AR_2 contains a nonzero nilpotent element of index two; that is, there exists a matrix

$$X = \begin{vmatrix} x & y \\ v & w \end{vmatrix}$$

such that $AX \neq 0$ but $(AX)^2 = 0$. This gives rise to (1).

Conversely, let M be any nonzero matrix of R_2. By multiplying M by the appropriate E_{i1}, we can show that MR_2 contains a nonzero matrix of the form

$$A = \begin{vmatrix} a & 0 \\ b & 0 \end{vmatrix}.$$

Now (1) implies that there exists a matrix

$$X = \begin{vmatrix} x & y \\ 0 & 0 \end{vmatrix}$$

such that $AX \neq 0$ but $(AX)^2 = 0$. Hence R_2 is DN.

PROPOSITION 5. *If R is commutative, then R_n $(n > 1)$ is DN.*

Proof. Assume $n = 2$ and use the notation of Lemma 4. If $a^2 \neq 0$, let $x = b$ and $y = -a$. Then (1) is satisfied. If $a^2 = 0$, then either $ba = 0$, in which case we let $x = 1$ and $y = 0$, or $ba \neq 0$, in which case we let $x = a$ and $y = 0$. Then (1) is satisfied. Thus, by Lemma 3, R_n is DN.

Recall that a domain R is a left Ore domain if for nonzero elements a, b of R there exist nonzero elements x, y of R such that $xa = yb$. For example, left Noetherian domains and commutative domains are left Ore domains. Among domains, left Ore domains are precisely those which have a classical ring of quotients.

PROPOSITION 6. *Let R be a domain. Then R is a left Ore domain if and only if R_n $(n > 1)$ is DN.*

Proof. Let R be a left Ore domain, and let M be a nonzero matrix of R_2. Then MR_2 contains a nonzero matrix A of the form

$$\begin{vmatrix} a & 0 \\ b & 0 \end{vmatrix}.$$

When $a = 0$ then $A^2 = 0$, and when $b = 0$, then $AE_{12} \neq 0$ but $(AE_{12})^2 = 0$. Now consider the case in which $a \neq 0$ and $b \neq 0$. By the left Ore condition, there exist $x, y \in R$ such that $xa = (-y)b \neq 0$. Let $X = xE_{11} + yE_{12}$. Then $AX \neq 0$ and each entry in $(AX)^2$ has $xa + yb$ as a factor. Hence $(AX)^2 = 0$. Therefore, from Lemma 3, R_n is DN.

Conversely, assume that R_n $(n > 1)$ is DN and there exists an infinite direct sum $\bigoplus_{t \in T} H_t$ of nonzero left ideals of R. Let A be the $n \times n$ matrix whose first column consists of nonzero elements of R each from a different H_t and whose other columns consist of zeros. Since R_n is DN there exists a $n \times n$ matrix X such that $AX \neq 0$ but $(AX)^2 = 0$. The (i, j)-entry of $(AX)^2$ is $a_i(\Sigma_{k=1}^n x_k a_k)x_j$ where a_r

is the $(r, 1)$ entry in A and x_s is the $(1, s)$ entry in X. However, R is a domain; thus $\Sigma_{k=1}^{n} x_k a_k = 0$ which contradicts the assumption that $\bigoplus_{t \in T} H_t$ is an infinite direct sum of nonzero left ideals. Hence R has finite uniform (Goldie) dimension which implies R is a left Ore domain [3; pp. 343–345].

We wish to thank the referee for his suggestions, which greatly improved the exposition of this paper.

References

1. F. W. Anderson and K. R. Fuller, Rings and Categories of Modules, Springer–Verlag, New York-Heidelberg-Berlin, 1973.

2. G. F. Birkenmeier, Self-injective rings and the minimal direct summand containing the nilpotents, Communications in Algebra, 4(8) (1976) 705–721.

3. C. Faith, Algebra: Rings, Modules and Categories I, Springer–Verlag, New York-Heidelberg-Berlin, 1973.

Division of Natural Sciences and Mathematics, Floyd Junior College, P. O. Box, 1864, Rome, GA 30161.

Department of Mathematics, University of Wisconsin–Milwaukee, Milwaukee, WI 53201.

CORRECTION TO "A CHARACTERIZATION OF INTEGERS"

(This Monthly, 84 (1977) 278–281)

Maurice Mignotte

On page 279, in line 2 and line 6, after "$\|\lambda\theta^n\| \leqq \varepsilon$" and "$\|\lambda\theta^n\| \leqq 10\varepsilon$," respectively, the following should be added:

$$\text{with} \quad \varepsilon = [2e\delta(\delta+1)\ (1 + \operatorname{Log}\lambda)]^{-1}.$$

Université Louis Pasteur, Centre de Calcul, 7 Rue René Descartes, 67084 Strasbourg, France.

RESEARCH PROBLEMS

Edited by Richard Guy

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics and Statistics, The University of Calgary, Calgary, Alberta, Canada, T2N 1N4.

HOW MANY 3-TERM ARITHMETIC PROGRESSIONS CAN THERE BE IF THERE ARE NO LONGER ONES?

G. J. Simmons and H. L. Abbott

Introduction. Let $\mathscr{A}_k(n)$ be the set of n-term nonnegative integer sequences which contain no k-term A.P. as a subsequence. Denote by $f(A)$ the number of 3-term A.P.'s in the sequence A. Erdös [1] defined

$$f_k(n) = \max_{A \in \mathscr{A}_k(n)} f(A)$$

and showed that $\log f_4(n)/\log n > 1.4649$ infinitely often by exhibiting sequences $A \in \mathcal{A}_4(3^r)$ for which $f(A) = 5^{r-1}$. Erdös and Simmons proved [1] that for all $k > 3$

$$\lim_{n\to\infty} \frac{\log f_k(n)}{\log n} = s_k$$

exists, but were unable to show $s_k < 2$. Erdös has conjectured that $f_4(n) = o(n^2)$ and perhaps that $f_4(n) < n^{2-\varepsilon}$. In this note we show that $f_4(n) \geqq n^{1.623}$ infinitely often and that $s_k \to 2$ as $k \to \infty$ and reopen the question of determining $\lim_{n\to\infty} f_4(n)$, and more generally, of determining s_k.

The replication recursion. Let $A = (0, a_1, a_2, \ldots, a_{m-1})$ and $B = (0, b_1, b_2, \ldots, b_{n-1})$ be a pair of sequences having no k-term A.P.'s, $k > 3$, but having $g(m)$ and $g(n)$ 3-term A.P.'s respectively. Assume with no loss of generality that $a_{m-1} \geqq b_{n-1}$. Define $u = 2a_{m-1} + 1$ and construct an mn term sequence C

$$C = \sum_{i=0}^{n-1} (b_i u + A) \tag{1}$$

by replicating A n times with the indicated translations $b_i u$. Clearly no A.P. in C, which has two terms in one of the replications of A (blocks), can have a third term in another block, i.e., every 3-term A.P. is either wholly in one block or else each term is in a distinct block. It is a simple matter therefore to enumerate all of the possible 3-term A.P.'s in C. There are a total of $ng(m)$ 3-term A.P.'s each of whose terms are all in the same block. The number of blocks in arithmetic progression is $g(n)$. The i^{th} term of each of these blocks is by construction a term in an A.P., which contributes $mg(n)$ additional 3-term A.P.'s.

Finally, given the corresponding 3-term A.P.'s in three blocks which are in arithmetic progression themselves, two new 3-term A.P.'s are possible:

$$(a, a+d, a+2d) \qquad (a+\delta, a+\delta+d, a+\delta+2d) \qquad (a+2\delta, a+2\delta+d, a+2\delta+2d)$$

which contributes $2g(m)g(n)$ additional A.P.'s. The sequence C therefore has by construction $mg(n) + ng(m) + 2g(m)g(n)$ 3-term A.P.'s, hence

$$f_k(mn) \geqq f(C) = g(mn) = mg(n) + ng(m) + cg(m)g(n), \tag{2}$$

where $c = 2$.

Functional equations of the form of (2) arise frequently in situations where two constructions can be compounded to form a new construction by indexing replications of one by the elements of the other. In the general case, the constant multiplier, c, in the righthand term may take any non-zero value. The general solution to equation (2) is

$$g(n) = (n^s - n)/c, \qquad -\infty \leqq s < \infty. \tag{3}$$

where $g(1) \neq 0$ implies $s = -\infty$ and $g(n) = -(n/c)$, and $g(0) \neq 0$ implies $s = 0$ and $g(n) = (1-n)/c$. Without further information, the solution to (2) given by (3) does not bound $f_k(n)$.

Given a sequence $A \in \mathcal{A}_k(p)$ containing $f(A) = g_k(p)$ 3-term A.P.'s the constructed lower bound for $f_k(p^t)$ is given by

$$p^t + 2g_k(p^t) = (p + 2g_k(p))^t. \tag{4}$$

By combining equations (3) and (4), we obtain

$$s_k \geqq \frac{\log(p + 2g_k(p))}{\log p} \tag{5}$$

since this value will be realized for $p, p^2, \ldots$ term sequences using the construction described above.

Since in a sequence of $k-1$ consecutive integers there are $[(k-1)(k-3)]/4$ 3-term A.P.'s if k is odd or $(k-2)^2/4$ 3-term A.P.'s if k is even,

$$g_k(k-1) \geqq \begin{cases} \dfrac{(k-1)(k-3)}{4} & k \text{ odd} \\ \dfrac{(k-2)^2}{4} & k \text{ even.} \end{cases} \tag{6}$$

Substituting the lower bounds in (6) into (5) yields

$$s_k \geqq \begin{cases} 2-\log 2/\log(k-1) & k \text{ odd} \\ \dfrac{\log(k^2-2k+2)-\log 2}{\log(k-1)} & k \text{ even} \end{cases} \tag{7}$$

and hence the result that $\lim_{k\to\infty} s_k = 2$.

Numerical results for $k = 4$. Clearly $f_4(1)=f_4(2)=0$, $f_4(3)=1$, $f_4(4)=2$ and $f_4(5)=3$. Representative best constructions known to date for the next few values of n and the associated values of $g(n)$ are $0,1,2,4,5,8$ which gives $f_4(6)\geqq 4$; $0,1,2,4,7,8,14$ which gives $f_4(7)\geqq 6$; $0,1,2,4,5,7,8,9$ which gives $f_4(8)\geqq 8$; $0,2,4,8,9,10,14,16,18$ which gives $f_4(9)\geqq 12$; $0,2,4,8,9,10,14,16,18,28$ which gives $f_4(10)\geqq 15$ and the translation by 19 of the eleven-term sequence $-19,-9,-7,-5,-1,0,1,5,7,9,19$ which gives $f_4(11)\geqq 19$. Using the construction for which $g(11)=19$, inequality (5) gives

$$s_4 \geqq \frac{\log 49}{\log 11} \approx 1.623\ldots.$$

Conclusion. We have considered the following alternate approach to the problem, which suggests that $s_k = 2$ for all $k \geqq 4$. Denote by $r_k(x)$ the size of a maximal subset of $\{1,2,\ldots,x\}$ containing no k-term arithmetic progression. Let $n = r_4(x)$ and let $A = \{a_1, a_2, \ldots, a_n\}$ be a subset of $\{1,2,\ldots,x\}$ which contains no 4-term arithmetic progression. We have been able to show, using an argument that parallels closely that used in a paper of P. Varnavides [2], that if for every ε, $0<\varepsilon<1$, the estimate

$$\frac{r_3(x^\varepsilon)}{x^\varepsilon} = o\left(\frac{r_4(x)}{x}\right) \tag{8}$$

holds, then A contains at least $n^{2-\varepsilon}$ 3-term arithmetic progressions.

At the present time we do not have enough information about the behavior of $r_k(x)$ to decide (8), in spite of E. Szemerédi's recent remarkable work on the problem [3]. However, almost any reasonable conjecture as to the behavior of $r_k(x)$ implies (8). The reader should consult [3] for further references.

This work was supported in part by the U.S. Energy Research & Development Administration (ERDA) under Contract No. AT(29–1)–789.

References

1. P. Erdös, Problems and results on combinatorial number theory, Chapter 12 (esp. p. 119) in *A Survey of Combinatorial Theory* edited by J. Srivastava, North Holland, 1973.

2. P. Varnavides, On certain sets of positive density, J. London Math. Soc., 34 (1959) 358–360.

3. E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, Acta Arith., XXVII (1975) 199–245.

Sandia Laboratories, Applied Mathematics Department, Albuquerque, NM 87115.
Department of Mathematics, University of Alberta, Edmonton, Canada T6G 2G1.

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Department of Mathematics, University of Wisconsin, Madison, WI 53706.

SOME REMARKS CONCERNING A DEFINITION OF ORDERED PAIRS

HUBERT H. SCHNEIDER

In any formal development of set theory, the concept of ordered pair plays a basic role; it is customary to define relations as classes of ordered pairs and functions as special relations. Using the notation $\langle A, B\rangle$ to indicate the ordered pair with first coordinate A and second coordinate B, ordered pairs are characterized by the property that $\langle A_1, B_1\rangle = \langle A_2, B_2\rangle$ iff $A_1 = A_2$ and $B_1 = B_2$. It is this characteristic property for ordered pairs that is used in a formal development of set theory; the ordered pair itself could be regarded as an undefined entity as is done in Bourbaki [**1**, p. 72]. On the other hand, it is possible to *define* ordered pairs in terms of basic class operations and thereby reducing the number of undefined terms. In 1914, Norbert Wiener [**6**] represented the ordered pair $\langle a, b\rangle$ of two sets a and b as the set $\{\{\{a\}\} \cup \{\varnothing\}\} \cup \{\{\{b\}\}\}$, where $\varnothing$ denotes the null set. At about the same time, Felix Hausdorff [**2**, p. 32] gave the definition of $\langle a, b\rangle$ as the set $\{\{a, 1\}, \{b, 2\}\}$, where 1 and 2 are two distinct objects different from a and b. In 1921, Kazimierz Kuratowski [**3**] offered a somewhat simpler definition for $\langle a, b\rangle$ that has been widely used ever since, namely the set $\{\{a, b\}, \{a\}\}$. These and other similar definitions for the ordered pair $\langle a, b\rangle$ are based on the assumption that its coordinates a and b are *sets*. This is adequate in set theories of the Zermelo–Fraenkel type where there are no proper classes (i.e., classes which are not sets). However, in set theories of the von Neumann–Bernays–Gödel type the above definitions are limited since they do not satisfy the characteristic property in case the coordinates are proper classes. On the other hand, there are many instances where one considers ordered pairs of proper classes, e.g., in the theory of categories. Below we shall discuss a definition of ordered pair, applicable to sets as well as proper classes, which is motivated by a suggestion of Quine [**5**, p. 202]. This definition, which of course has the characteristic property for ordered pairs, is such that the ordered pair is a set exactly in case both of its coordinates are sets. In addition, our definition leads to some simple relationships concerning the intersection, union and difference of ordered pairs.

For our discussions we shall assume the basic axioms of a von Neumann–Bernays–Gödel type set theory as developed in J. D. Monk [**4**]. Denoting by U the universal class, a class A is said to be a *set* iff $A \in U$. If $\phi(x)$ is a set-theoretical formula containing the free set variable x, we indicate by $\{x : \phi(x)\}$ the unique class of all sets x satisfying $\phi(x)$. Again, if $\tau(x)$ indicates a set-theoretical term in x, $\{\tau(x): \phi(x)\}$ denotes the class $\{y : y = \tau(x)$ and $\phi(x)$, for some set $x\}$. We note that $\{\tau(x): x \in A\} = \varnothing$ iff A is the null set $\varnothing$. If A is any class, the singleton $\{A\}$ is the class of all sets x such that $x = A$ whenever x is a set, i.e., $\{A\} = \{x : x = A$ if $A \in U\}$. We remark that $\{A\} = U$ iff A is a proper class. If f is any map and A any class, we denote by $f[A]$ the f-image of A, i.e., $f[A] = \{f(x): x \in A\}$.

Our definition of the ordered pair of two classes is based on two maps from the universal class U to itself, as follows. Let f_1 and f_2 be two maps from U into U which are such that

(*) f_1 and f_2 are one-one maps with disjoint ranges.

Then the ordered pair $\langle A, B\rangle$, having the class A as its first coordinate and the class B as its second coordinate, is defined by:

$$\langle A, B\rangle = f_1[A] \cup f_2[B].$$

There are infinitely many maps from U to U satisfying the condition (*). For example, consider the maps f_1 and f_2 such that for all $x \in U$: $f_1(x) = \{\{x\}\}$ and $f_2(x) = \{x, \varnothing\}$. f_1 and f_2 are obviously one-one maps with disjoint ranges; thus for these maps our definition of ordered pair yields:

$$\langle A, B \rangle = \{\{\{x\}\}: x \in A\} \cup \{\{x, \varnothing\}: x \in B\}.$$

Our first theorem which shows that the above definition of ordered pair is indeed appropriate, is a consequence of condition (*).

THEOREM 1. *$\langle A, B \rangle = \langle C, D \rangle$ iff $A = C$ and $B = D$.*

Proof. Suppose $\langle A, B \rangle = \langle C, D \rangle$ so that $f_1[A] \cup f_2[B] = f_1[C] \cup f_2[D]$. Since f_1 and f_2 are maps with disjoint ranges, it follows that $f_1[A] = f_1[C]$ and $f_2[B] = f_2[D]$. Again, since f_1 and f_2 are one-one maps, we get $A = C$ and $B = D$. Conversely, if $A = C$ and $B = D$, then we have trivially $\langle A, B \rangle = \langle C, D \rangle$.

We note that Theorem 1 is in fact equivalent with condition (*). Indeed, suppose that $f_1(x) = f_1(y)$ and hence $f_1[\{x\}] = f_1[\{y\}]$. Then $\langle \{x\}, \varnothing \rangle = f_1[\{x\}] = f_1[\{y\}] = \langle \{y\}, \varnothing \rangle$ and hence, assuming Theorem 1, $\{x\} = \{y\}$. Thus $x = y$ which shows that f_1 is a one-one map. Analogously it follows that f_2 is one-one. Next, suppose $f_1(x) = f_2(y)$ for some sets x and y; then $\langle \{x\}, \varnothing \rangle = \langle \varnothing, \{y\} \rangle$ so that on account of Theorem 1 we would get $\{x\} = \varnothing = \{y\}$ which is impossible. Hence f_1 and f_2 have disjoint ranges.

In addition to the characteristic property for ordered pairs, as expressed in Theorem 1, it is obviously desirable that the ordered pair of two sets is itself a set. The next theorem shows that our definition for the ordered pair has this property.

THEOREM 2. *$\langle A, B \rangle$ is a set iff both A and B are sets.*

Proof. f_1 and f_2 are one-one maps so that by the axiom of substitution: $A \in U$ iff $f_1[A] \in U$, and $B \in U$ iff $f_2[B] \in U$. Hence, if $A \in U$ and $B \in U$ then $\langle A, B \rangle = f_1[A] \cup f_2[B] \in U$ by the union axiom. Conversely, if $\langle A, B \rangle = f_1[A] \cup f_2[B] \in U$, then $f_1[A] \in U$ and $f_2[B] \in U$ by the subset axiom, and thus $A \in U$ as well as $B \in U$.

Since f_1 and f_2 are maps with *disjoint* ranges, it follows by simple truth-functional argument that for any classes A, B, C, and D:

$$(f_1[A] \cup f_2[B]) * (f_1[C] \cup f_2[D]) = (f_1[A] * f_1[C]) \cup (f_2[B] * f_2[D]),$$

where $*$ denotes any of the binary class operations $\cup$, $\cap$, or $-$. Furthermore, since f_1 and f_2 are *one-one* maps, it is easily seen that for any classes A, B, C, and D: $f_1[A] * f_1[C] = f_1[A * C]$ and $f_2[B] * f_2[D] = f_2[B * D]$. In view of these identities we have thus the following relationships concerning certain algebraic operations between ordered pairs and their coordinates.

THEOREM 3. (1) $\langle A, B \rangle \cup \langle C, D \rangle = \langle A \cup C, B \cup D \rangle$,

(2) $\langle A, B \rangle \cap \langle C, D \rangle = \langle A \cap C, B \cap D \rangle$,

(3) $\langle A, B \rangle - \langle C, D \rangle = \langle A - C, B - D \rangle$.

As immediate consequences of these properties we obtain a corresponding identity concerning the operation $\oplus$ of symmetric difference applied to ordered pairs, and an equivalence regarding the inclusion relation between ordered pairs and their coordinates.

COROLLARY. (4) $\langle A, B \rangle \oplus \langle C, D \rangle = \langle A \oplus C, B \oplus D \rangle$,

(5) *$\langle A, B \rangle \subseteq \langle C, D \rangle$ iff $A \subseteq C$ and $B \subseteq D$.*

Proof of (5): $\langle A, B \rangle \subseteq \langle C, D \rangle$ iff $\langle A, B \rangle \cap \langle C, D \rangle = \langle A, B \rangle$,

iff $\langle A \cap C, B \cap D \rangle = \langle A, B \rangle$ by (2),

iff $A \cap C = A$ and $B \cap D = B$ by Theorem 1,

iff $A \subseteq C$ and $B \subseteq D$.

If $\bar{A} = U - A$ and $\bar{B} = U - B$ are the complements of A and B, respectively, and we define the relative complement $\widetilde{\langle A, B\rangle}$ of the pair $\langle A, B\rangle$ by $\widetilde{\langle A, B\rangle} = \langle U, U\rangle - \langle A, B\rangle$, Theorem 3(3) yields at once the following property.

COROLLARY. (6) $\widetilde{\langle A, B\rangle} = \langle \bar{A}, \bar{B}\rangle$.

On account of Theorem 3 and this last corollary it is apparent that the collection of ordered pairs satisfies the laws for a Boolean lattice having the ordered pair $\langle U, U\rangle$ as its unit and the ordered pair $\langle \varnothing, \varnothing\rangle$ as its zero. Again, if S_1 and S_2 are any fixed classes, then the collection of ordered pairs $\langle A, B\rangle$ with $A \subseteq S_1$ and $B \subseteq S_2$ forms a Boolean lattice where now the ordered pair $\langle S_1, S_2\rangle$ is the unit and $\widetilde{\langle A, B\rangle} = \langle S_1 - A, S_2 - B\rangle$ is the complement of $\langle A, B\rangle$. In the special case where S_1 is a set with m elements and S_2 is a set with n elements, this Boolean lattice has 2^{m+n} elements.

Considering the collection of ordered pairs $\langle A, B\rangle$ such that $A \subseteq B \subseteq S$ for some fixed class S, we note that this collection is closed with respect to the operations of union and intersection; indeed, if $A \subseteq B$ and $C \subseteq D$ then $A \cup C \subseteq B \cup D$ and $A \cap C \subseteq B \cap D$; hence in view of Theorem 3(1) and (2) we find that $\langle A, B\rangle \cup \langle C, D\rangle$ and $\langle A, B\rangle \cap \langle C, D\rangle$ belong to this collection whenever $\langle A, B\rangle$ and $\langle C, D\rangle$ belong to it. Thus the collection of ordered pairs $\langle A, B\rangle$ with $A \subseteq B \subseteq S$ forms a distributive lattice with unit $\langle S, S\rangle$ and zero $\langle \varnothing, \varnothing\rangle$.

It should be pointed out that one minor defect in our approach to defining the ordered pair is that it only makes sense in a set theory in which all objects are sets or classes. The standard definition of ordered pair does not have this limitation and may thus be presented to students before they understand that all mathematics is a branch of set theory.

References

1. N. Bourbaki, Theory of Sets, Addison–Wesley, Reading, Mass. 1968.
2. F. Hausdorff, Grundzüge der Mengenlehre, Veit, Leipzig, 1914.
3. K. Kuratowski, Sur la notion d'ordre dans la théorie des ensembles, Fund. Math., 2 (1921) 161–171.
4. J. D. Monk, Introduction to Set Theory, McGraw–Hill, New York, 1969.
5. W. V. Quine, Mathematical Logic, Harvard University Press, Cambridge, 1958.
6. N. Wiener, A Simplification of the Logic of Relations, Proc. Camb. Phil. Soc., 17 (1912–14) 387–390.

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF NEBRASKA, LINCOLN, NE 68508.

AN ELEMENTARY APPROACH TO SOME QUALITATIVE RESULTS IN ORDINARY DIFFERENTIAL EQUATIONS

JOHN V. BAXLEY

One of the earliest and most celebrated results in the qualitative theory of ordinary differential equations is the 1831 comparison theorem of Sturm. Since its proof involves no more than a simple use of integration by parts, this theorem occurs in several widely used textbooks, for example [**4**, p. 117, Ex. 4] and [**7**, p. 122]. One important use of Sturm's theorem is to show that solutions of certain equations (for example, Bessel's equation [**4**, p. 179, Ex. 6]) oscillate infinitely often as $x \to +\infty$.

It is the purpose of this note to discuss and apply a different comparison theorem, due to Čaplygin [**1**, p. 139], complementary to that of Sturm in that it may be used to investigate the qualitative behavior of non-oscillatory solutions of ordinary differential equations. An amusing contrast is that the proof of this theorem requires only a simple application, not of integration by parts, but of the second derivative test. The author makes no claim to originality of facts or methods here; they are well

known to the experts, particularly to those acquainted with the maximum principle [**6**, Chapter 1]. The elementary presentation given here, together with the application to Airy's equation, seems to the author to merit wider attention. We use only the techniques of elementary calculus and basic facts about differential equations contained in almost any elementary textbook (see [**2**], [**3**], [**4**], or [**7**]).

To illustrate the flair with which the second derivative test can be used in a particular transparent example, consider Legendre's equation

$$(1-x^2)y''-2xy'+n(n+1)y=0, \tag{1}$$

where $n \geqq 0$ is an integer. The n^{th} Legendre polynomial $P_n(x)$ is a solution of (1) for all x. It is a fact that $P_n(x)$ has n distinct zeros, all real, and all lying in the interval $-1<x<1$. We content ourselves here with verifying only that $P_n(x)$ has no real zeros outside $-1<x<1$. Since $P_n(x)$ is an even function for n even, and an odd function for n odd, it suffices to show that $P_n(x)>0$ for $x \geqq 1$. Actually, we shall show that $P_n'(x)>0$ for $x \geqq 1$ and since $P_n(1)=1$, it follows that $P_n(x)$ is increasing on $1 \leqq x < \infty$, so $P_n(x) \geqq 1$ for $x \geqq 1$. Clearly we may assume $n \geqq 1$.

From (1), we have

$$(1-x^2)P_n''(x)-2xP_n'(x)+n(n+1)P_n(x)=0 \tag{2}$$

for all x. Letting $x=1$ gives $2P_n'(1)=n(n+1)$. Thus $P_n'(1)>0$ for $n \geqq 1$. Suppose $P_n'(x)>0$ for $x \geqq 1$ is false. Then there exists $C>1$ for which $P_n'(C)=0$ and we may, without loss of generality, assume $P_n'(x)>0$ for $1<x<C$, and hence that $P_n(x)$ is increasing on $1<x<C$. Then $P_n(C)>P_n(1)=1$ and putting $x=C$ in (2), we get

$$(C^2-1)P_n''(C)=n(n+1)P_n(C)>0,$$

and so $P_n''(C)>0$. Applying the second derivative test, we conclude that $P_n(x)$ has a minimum at $x=C$, contradicting that $P_n(x)$ is increasing on $1<x<C$. Thus $P_n'(x)>0$ for $x \geqq 1$.

With this background, we now pass to the promised comparison theorem, whose proof will be seen as a replay of the above argument.

COMPARISON THEOREM. *Let $L(y)=y''+p(x)y'+q(x)y$, where $q(x)<0$ on an interval I. Suppose that for some point $a \in I$, $u(a) \geqq v(a)$, $u'(a)>v'(a)$, and further that $L(u(x)) \geqq L(v(x))$ on I. Then $u'(x)>v'(x)$, $u(x)>v(x)$ for all $x \in I$ such that $x>a$.*

Proof. Let $w(x)=u(x)-v(x)$. It suffices to show that $w'(x)>0$ for $x \in I$, $x>a$, for then $w(x)>0$ for $x \in I$, $x>a$, and the desired conclusion is immediate. If we assume the contrary, then there exists $c \in I$, $c>a$, for which $w'(c)=0$ and, without loss of generality, we may assume $w'(x)>0$ for $a<x<c$. Thus $w(c)>w(a) \geqq 0$, and putting $x=c$ into $L(w(x)) \geqq 0$ gives $w''(c) \geqq -q(c)w(c)>0$ and the second derivative test implies a minimum for $w(x)$ at c, which as before contradicts $w'(x)>0$ for $a<x<c$.

We shall use this simple theorem to get rather precise knowledge about the behavior of solutions to Airy's equation

$$y''-xy=0 \tag{3}$$

as $x \to +\infty$. A naive sophomore might make the mistake of "solving" (3) by treating it as a constant coefficient equation and thus mistakenly believe that $\exp(\sqrt{x}\,x)=\exp(x^{\frac{3}{2}})$ and $\exp(-\sqrt{x}\,x)=\exp(-x^{\frac{3}{2}})$ are linearly independent solutions of (3). Actually such a conclusion is not completely preposterous; for rather sharp estimates on the solutions of Airy's equation are known and are obtained in [**5**, p. 138] from results on the asymptotic behavior of Bessel functions. These results state essentially that Airy's equation has two linearly independent solutions, one of which grows like $x^{-\frac{1}{4}}\exp(\frac{2}{3}x^{\frac{3}{2}})$ as $x \to +\infty$ and the other of which decays like $x^{-\frac{1}{4}}\exp(-\frac{2}{3}x^{\frac{3}{2}})$. It is too much to expect that our elementary comparison theorem could give this sharp result, but we shall see that our theorem can come within epsilon of this result.

Let $V_q(x) = x^{-q}\exp(\frac{2}{3}x^{\frac{3}{2}})$, where $0 < q < 1$. We shall prove

THEOREM. *Let r be fixed, $\frac{1}{4} < r < 1$. Then there exist two linearly independent solutions $\phi_1(x)$ and $\phi(x)$ of Airy's equation and $x_r > 1$ such that*

$$V_r(x) < \phi_1(x) < V_{\frac{1}{4}}(x), \qquad x_r < x < +\infty,$$

$$\lim_{x\to+\infty} x^p V_r(x)\phi(x) = 0, \qquad \textit{if} \quad 0 < p < \tfrac{1}{2},$$

$$\lim_{x\to+\infty} x^p V_{\frac{1}{4}}(x)\phi(x) = +\infty, \qquad \textit{if} \quad p > \tfrac{1}{2}.$$

At first glance, it might seem that applying the theorem for, say, $r = \frac{3}{8}$ would give solutions whose asymptotic behavior is perhaps different and more tightly bounded than the solutions obtained with $r = \frac{1}{2}$. However, each of the solutions obtained using $r = \frac{1}{2}$ is a linear combination of the solutions obtained using $r = \frac{3}{8}$, and therefore asymptotically behave the same. Using this idea, we easily prove

COROLLARY. *Airy's equation has two linearly independent solutions $\phi_1(x)$ and $\phi(x)$ for which*

$$\phi_1(x) < V_{\frac{1}{4}}(x), \qquad \textit{for } x \textit{ sufficiently large},$$

$$\lim_{x\to+\infty} \frac{\phi_1(x)}{V_q(x)} = +\infty, \qquad \textit{if} \quad \tfrac{1}{4} < q < 1,$$

$$\lim_{x\to+\infty} x^p V_q(x)\phi(x) = 0, \qquad \textit{if} \quad \tfrac{1}{4} < q < 1,\ 0 < p < \tfrac{1}{2},$$

$$\lim_{x\to+\infty} x^p V_{\frac{1}{4}}(x)\phi(x) = +\infty, \quad \textit{if} \quad p > \tfrac{1}{2}.$$

We now prove the theorem. An easy calculation gives

$$V_q''(x) - xV_q(x) = x^{-q-\frac{1}{2}}\exp(\tfrac{2}{3}x^{\frac{3}{2}})\left[\frac{q(q+1)}{x^{\frac{3}{2}}} - (2q - \tfrac{1}{2})\right].$$

Thus, we see that $V_{\frac{1}{4}}''(x) - xV_{\frac{1}{4}}(x) > 0$ for $x > 0$ and since $\frac{1}{4} < r < 1$, there exists $x_r > 1$ such that $V_r''(x) - xV_r(x) < 0$ if $x \geqq x_r$. Since $x_r > 1$ and $\frac{1}{4} < r < 1$, we easily see that

$$V_r(x_r) < V_{\frac{1}{4}}(x_r) \qquad V_r'(x_r) < V_{\frac{1}{4}}'(x_r).$$

Choose numbers y_0 and y_1 satisfying

$$V_r(x_r) < y_0 < V_{\frac{1}{4}}(x_r)$$
$$V_r'(x_r) < y_1 < V_{\frac{1}{4}}'(x_r).$$

Let $\phi_1(x)$ be the unique solution of (3) which satisfies the initial conditions $y(x_r) = y_0$, $y'(x_r) = y_1$. (We use here the existence and uniqueness theorems.) Then, according to our comparison theorem,

$$V_r(x) < \phi_1(x) < V_{\frac{1}{4}}(x)$$
$$V_r'(x) < \phi_1'(x) < V_{\frac{1}{4}}'(x), \qquad \text{for} \quad x > x_r.$$

One might expect that the exponentially damped solution $\phi(x)$ could be found in a similar fashion. However, a little reflection shows this is not the case. As we shall see shortly, the comparison method will construct a solution $\phi_2(x)$, linearly independent of $\phi_1(x)$, which behaves as $x \to +\infty$ like $\phi_1(x)$. Therefore, if this same method could construct an exponentially damped solution, it would construct two linearly independent exponentially damped solutions from which it would follow that all solutions

are exponentially damped. Our plan is to construct two linearly independent exponentially growing solutions $\phi_1(x)$ and $\phi_2(x)$ and search for a particular linear combination which is exponentially damped. For this purpose we choose a number $\bar{y}_1 \neq y_1$ which still satisfies $V_r'(x_r) < \bar{y}_1 < V_{\frac{1}{4}}'(x_r)$. Repeating the above argument with y_1 replaced by $\bar{y}_1$, we get a solution $\phi_2(x)$ of (3) for which

$$\begin{aligned} &V_r(x) < \phi_2(x) < V_{\frac{1}{4}}(x) \\ &V_r'(x) < \phi_2'(x) < V_{\frac{1}{4}}'(x), \qquad \text{for} \quad x > x_r. \end{aligned}$$

However, $\phi_1(x)$ and $\phi_2(x)$ are linearly independent since the Wronskian of $\phi_1(x)$ and $\phi_2(x)$ evaluated at x_r is

$$W(\phi_1, \phi_2)(x_r) = y_0(y_1 - \bar{y}_1) \neq 0.$$

Note that since Airy's equation has no y' term, by Abel's formula for the Wronskian, $W(\phi_1, \phi_2)(x)$ is a constant C.

Now every solution $\phi(x)$ of Airy's equation can be expressed in the form

$$\phi(x) = C_1\phi_1(x) + C_2\phi_2(x).$$

We show next that it is possible to choose C_1 and C_2 so that $\phi(x) \to 0$ as $x \to +\infty$. Write

$$\phi(x) = \phi_1(x)\left[C_1 + C_2\frac{\phi_2(x)}{\phi_1(x)}\right], \quad \text{for} \quad x \geqq x_r.$$

Then by the Fundamental Theorem of Calculus,

$$\frac{\phi_2(x)}{\phi_1(x)} = \frac{\phi_2(x_r)}{\phi_1(x_r)} + \int_{x_r}^{x}\left[\frac{\phi_2(t)}{\phi_1(t)}\right]' dt, \quad \text{for} \quad x \geqq x_r.$$

Since

$$\left[\frac{\phi_2(t)}{\phi_1(t)}\right]' = \frac{W(\phi_1, \phi_2)(x)}{\phi_1^2(t)} = \frac{C}{\phi_1^2(t)}$$

and $\phi_1(t)$ is exponentially increasing, an easy comparison shows that the improper integral

$$\int_{x_r}^{\infty}\left[\frac{\phi_2(t)}{\phi_1(t)}\right]' dt$$

converges absolutely. Thus

$$\lim_{x\to\infty}\frac{\phi_2(x)}{\phi_1(x)} \equiv k$$

exists and is finite. Choosing $C_1 = -k$ and $C_2 = 1$,

$$\phi(x) = \phi_1(x)\left[\frac{\phi_2(x)}{\phi_1(x)} - k\right],$$

and by construction, the expression in brackets above tends to 0 as $x \to +\infty$. Thus there is hope that $\phi(x) \to 0$ as $x \to +\infty$. We actually calculate, using L'Hospital's rule (here $p > 0$, $0 < q < 1$)

$$\begin{aligned} \lim_{x\to+\infty} x^p V_q(x)\phi(x) &= \lim_{x\to+\infty}\left[\frac{\phi_2(x)}{\phi_1(x)} - k\right]\Big/\left[\frac{1}{x^p V_q(x)\phi_1(x)}\right] \\ &= \lim_{x\to+\infty} -Cx^p\Big/\left[\frac{p}{x}\frac{\phi_1(x)}{V_q(x)} + \frac{\phi_1(x)}{V_q(x)}\frac{V_q'(x)}{V_q(x)} + \frac{\phi_1'(x)}{V_q'(x)}\frac{V_q'(x)}{V_q(x)}\right]. \end{aligned}$$

Letting $q = r$, we see that the denominator in this last fraction is bounded below by

$$\frac{p}{x} + 2x^{\frac{1}{2}} - \frac{2r}{x}$$

and thus $\lim_{x\to\infty} x^p V_r(x)\phi(x) = 0$, if $p < \frac{1}{2}$. Letting $q = \frac{1}{4}$, we see that the same denominator is bounded above by

$$\frac{p}{x} + 2x^{\frac{1}{2}} - \frac{1}{2x}$$

and thus

$$\lim_{x\to+\infty} x^p V_{\frac{1}{4}}(x)\phi(x) = +\infty, \quad \text{if} \quad p > \tfrac{1}{2}$$

and the theorem is proved since the linear independence of $\phi_1(x)$ and $\phi_2(x)$ is obvious.

References

1. E. F. Beckenbach and R. Bellman, Inequalities, Springer-Verlag, Berlin, 1961.
2. W. E. Boyce and R. C. DiPrima, Introduction to Differential Equations, Wiley, New York, 1970.
3. Fred Brauer and J. E. Nohel, Ordinary Differential Equations: A First Course, Benjamin, Reading, Mass., 1973.
4. E. A. Coddington, An Introduction to Ordinary Differential Equations, Prentice-Hall, Englewood Cliffs, N. J., 1961.
5. N. N. Lebedev, Special Functions and Their Applications, Prentice-Hall, Englewood Cliffs., N. J., 1965.
6. M. H. Protter and H. F. Weinberger, Maximum Principles in Differential Equations, Prentice-Hall, Englewood Cliffs, N.J., 1967.
7. G. F. Simmons, Differential Equations, with Applications and Historical Notes, McGraw-Hill, New York, 1972.

Department of Mathematics, Wake Forest University, Winston-Salem, NC 27109.

GALOIS RESOLVENTS OF PERMUTATION GROUPS

Phyllis Lefton

In this note the author approaches Galois resolvents from a viewpoint more general than that taken in the literature ([**1**, III §5], [**2**, §4.8], [**4**, p. 216]). Here, a Galois resolvent is associated to a permutation group rather than to a specific polynomial. This general resolvent can then be specialized with respect to a given polynomial over $\boldsymbol{Q}$ to give information about the Galois group of the splitting field of that polynomial. Although the author makes no claims for the originality of the results, she has been unable to find this general exposition in the literature. It can make a nice addition to an abstract algebra course once the Fundamental Theorem of Galois Theory has been proved and permutation groups and symmetric functions have been studied.

Let $\boldsymbol{Q}(x) = \boldsymbol{Q}(x_1, \ldots, x_n)$ be the field of rational functions in n indeterminates. The symmetric group Σ_n acts faithfully as a group of automorphisms of $\boldsymbol{Q}(x)$ by permuting the variables. The field of fixed elements in this action is the subfield $\boldsymbol{Q}(y) = \boldsymbol{Q}(y_1, \ldots, y_n)$ of rational functions in $y_1, \ldots, y_n$ the elementary symmetric functions in $x_1, \ldots, x_n$. Thus Σ_n may be regarded as the Galois group of $\boldsymbol{Q}(x)$ over $\boldsymbol{Q}(y)$.

By the Fundamental Theorem of Galois Theory, to each subgroup Π of Σ_n there corresponds an intermediate field $\boldsymbol{Q}(x)^{\Pi}$, the fixed field of Π. Since this field has finite degree over $\boldsymbol{Q}(y)$, we have $\boldsymbol{Q}(x)^{\Pi} = \boldsymbol{Q}(y, F(x))$ for some rational function $F(x) = F(x_1, \ldots, x_n)$, not uniquely determined. We may take $F(x)$ to be a polynomial in $x_1, \ldots, x_n$.

DEFINITION. Let $\Phi(z,y)$ be the minimal polynomial for $F(\boldsymbol{x})$ over $\boldsymbol{Q}(y)$. We call $\Phi(z,y)$ the Galois resolvent of Π corresponding to $F(\boldsymbol{x})$.

The roots of $\Phi(z,y)$ are the conjugates of $F(\boldsymbol{x})$ over Σ_n. That is, over $\boldsymbol{Q}(y)$, we have

$$\Phi(z,y)=\prod_{\pi\in S}(z-\pi F(\boldsymbol{x})), \tag{1}$$

where π runs through a set S of coset representatives of Σ_n/Π and $\pi F(\boldsymbol{x})=F(\pi\boldsymbol{x})$ with $\pi\boldsymbol{x}=(x_{\pi 1},\ldots,x_{\pi n})$. From (1) it is easy to see that $\Phi(z,y)$ is also the Galois resolvent of Π^{π} corresponding to $\pi F(\boldsymbol{x})$, where Π^{π} is the subgroup conjugate to Π over Σ_n, namely, $\Pi^{\pi}=\pi\Pi\pi^{-1}$, for $\pi\in S$. It also follows from (1) and the Theorem on Symmetric Polynomials that the coefficients of $\Phi(z,y)$ are polynomials in $y_1,\ldots,y_n$.

This general Galois resolvent can be "specialized" in relation to a given polynomial $f(x)\in\boldsymbol{Q}[x]$ as follows. Let

$$f(x)=x^n+a_1x^{n-1}+\cdots+a_n \tag{2}$$

be a polynomial with rational coefficients and distinct roots $\alpha_1,\ldots,\alpha_n$. By substituting $\boldsymbol{a}=(-a_1,a_2,\ldots,(-1)^na_n)$ for $\boldsymbol{y}$ and $\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\ldots,\alpha_n)$ for $\boldsymbol{x}$ in (1), we obtain

$$\Phi(z,\boldsymbol{a})=\prod_{\pi\in S}(z-\pi F(\boldsymbol{\alpha})). \tag{3}$$

This specialized resolvent is a polynomial in z whose coefficients and roots are certain polynomials in the coefficients and roots, respectively, of $f(x)$. In particular, we note that the coefficients of $\Phi(z,\boldsymbol{a})$ are rational numbers.

$\Phi(z,\boldsymbol{a})$ can give information about the Galois group G_f of the splitting field of $f(x)$. We recall that, since G_f acts faithfully on the roots of $f(x)$, there exists an injective homomorphism Ψ from G_f to Σ_n given by $\Psi(\sigma)=\pi_\sigma$, where

$$\sigma(\alpha_i)=\alpha_j \Leftrightarrow \pi_\sigma(i)=j \tag{4}$$

for all $i,j=1,\ldots,n$. We write $G_f\to\Pi$ if G_f is isomorphic, in the sense of (4), to a subgroup of a permutation group Π. We have the following relationship between the Galois group G_f of the splitting field of $f(x)$, given as in (2), and the Galois resolvent $\Phi(z,y)$ of a given subgroup Π of Σ_n.

THEOREM.
(a) *If $G_f\to\Pi^{\pi}$ for some $\pi\in\Sigma_n$, then $\Phi(z,\boldsymbol{a})$ has a rational root.*
(b) *Conversely, if $\Phi(z,\boldsymbol{a})$ has a rational root and no multiple roots, then $G_f\to\Pi^{\pi}$ for some $\pi\in\Sigma_n$.*

Proof. (a) If $G_f\to\Pi$, then for each $\sigma\in G_f$, $\pi_\sigma F(\boldsymbol{x})=F(\boldsymbol{x})$. Therefore, by (4), $\sigma F(\boldsymbol{\alpha})=F(\boldsymbol{\alpha})$, showing $F(\boldsymbol{\alpha})$ rational. More generally, suppose $G_f\to\Pi^{\pi}$ for some $\pi\in\Sigma_n$. We may assume $\pi\in S$ since there exists $\pi'\in S$ such that $\pi F(\boldsymbol{x})=\pi' F(\boldsymbol{x})$ and $\Pi^{\pi}=\Pi^{\pi'}$. Since $\Phi(z,\boldsymbol{a})$ is also the Galois resolvent of Π^{π} corresponding to $F(\pi\boldsymbol{x})$, the above argument, with Π, π_σ, $\boldsymbol{x}$, and $\boldsymbol{\alpha}$ replaced by Π^{π}, π_σ^{π}, $\pi\boldsymbol{x}$, and $\pi\boldsymbol{\alpha}$, respectively, shows that $F(\pi\boldsymbol{\alpha})$ is rational.

(b) Conversely, if $F(\boldsymbol{\alpha})$ is rational, then for each $\sigma\in G_f$, $\sigma F(\boldsymbol{\alpha})=F(\boldsymbol{\alpha})$. If, for some $\sigma\in G_f$, $\pi_\sigma F(\boldsymbol{x})\neq F(\boldsymbol{x})$, then, since $\Phi(z,\boldsymbol{a})$ has distinct roots, $\pi_\sigma F(\boldsymbol{\alpha})\neq F(\boldsymbol{\alpha})$. Therefore, $\sigma F(\boldsymbol{\alpha})\neq F(\boldsymbol{\alpha})$, a contradiction. If the rational root of $\Phi(z,\boldsymbol{a})$ is $\pi F(\boldsymbol{\alpha})$ for some $\pi\in S$, then we replace Π, π, $\boldsymbol{x}$, and $\boldsymbol{\alpha}$ by Π^{π}, π_σ^{π}, $\pi\boldsymbol{x}$ and $\pi\boldsymbol{\alpha}$, respectively, in the argument and get $G_f\to\Pi^{\pi}$.

The following corollary can be left as an easy exercise.

COROLLARY. *If $G_f\to\Delta$, where*

$$\Delta=\bigcap_{\pi\in\Sigma_n}\Pi^{\pi},$$

then all of the roots of $\Phi(z,\boldsymbol{a})$ are rational. The converse holds if $\Phi(z,\boldsymbol{a})$ has no multiple roots.

Examples. Galois resolvents for subgroups of Σ_n can be computed explicitly once an appropriate $F(\boldsymbol{x})$ is found.

1. Let A_n be the alternating group and $F = F(\boldsymbol{x}) = \Pi_{i<j}(x_i - x_j)$. Then $\boldsymbol{Q}(\boldsymbol{x})^{A_n} = \boldsymbol{Q}(y, F(\boldsymbol{x}))$, $\Phi(z, y) = z^2 - F^2$, and $\Phi(z, \boldsymbol{a}) = z^2 - D_f$, where D_f is the discriminant of $f(x)$. If $a_i \in \boldsymbol{Z}$ for $i = 1, \ldots, n$, then the theorem gives the well-known result ([**3**, p. 91]) that $G_f \to A_n \Leftrightarrow G_f = b^2$ for some $b \in \boldsymbol{Z}$.

2. Let C_n be the cyclic group generated by the cycle $(1\,2 \cdots n)$. It can be shown that $\boldsymbol{Q}(\boldsymbol{x})^{C_n} = \boldsymbol{Q}(y, F(\boldsymbol{x}))$ for $F(\boldsymbol{x}) = x_1^2x_2 + x_2^2x_3 + \cdots + x_n^2x_1$. The general Galois resolvent of C_n corresponding to this $F(\boldsymbol{x})$ is thus given by (1).

3. Let D_n be the subgroup of Σ_n generated by $\sigma = (1\ 2 \cdots n)$ and

$$\tau = (1\ n)(2\ n-1)\cdots\left(\frac{n}{2}n - \left[\frac{n}{2}\right]\right).$$

It can be shown that $\boldsymbol{Q}(\boldsymbol{x})^{D_n} = \boldsymbol{Q}(y, F(\boldsymbol{x}))$ for $F(\boldsymbol{x}) = x_1x_2 + x_2x_3 + \cdots + x_nx_1$. For $n = 4$, the specialized Galois resolvent of D_4 corresponding to $F(\boldsymbol{x})$ can be computed explicitly:

$$\Phi(z, \boldsymbol{a}) = z^3 - 2a_2z^2 + (a_2^2 + a_1a_3 - 4a_4)z + (a_3^2 + a_1^2a_4 - a_1a_2a_3).$$

It is a so-called "cubic resolvent" of $f(x)$ ([**4**, §58]). By the theorem, $G_f \to D_4$ if and only if the cubic resolvent has a rational root. A similar result, obtained by a somewhat different approach, appears in [**1**, p. 252].

References

1. N. Chebotarev, Grundzüge der Galois'schen Theorie, Noordhoff, Groningen, 1950.
2. L. Gaal, Classical Galois Theory, Chelsea, New York, 1973.
3. N. Jacobson, Lectures in Abstract Algebra, Vol. 3, Van Nostrand, Princeton, 1964.
4. D. H. Littlewood, A University Algebra, Dover, New York, 1970.
5. B. L. van der Waerden, Modern Algebra, Vol. I, Ungar, New York, 1953.

DEPARTMENT OF MATHEMATICS, MANHATTANVILLE COLLEGE, PURCHASE, NY 10577.

MATHEMATICAL EDUCATION

EDITED BY PAUL T. MIELKE AND SHIRLEY HILL

Material for this Department should be sent to Paul T. Mielke, Department of Mathematics, Wabash College, Crawfordsville, IN 47933, or to Shirley Hill, Department of Mathematics, University of Missouri, Kansas City, MO 64110.

CALCULUS IN HIGH SCHOOL — AT WHAT COST?

D. H. SORGE AND G. H. WHEATLEY

All across the country there is a clear trend of decline in norm referenced test scores. An article in the Chronicle of Higher Education (Scully, Malcolm G. Drop in Aptitude-Test Scores is Largest on Record. *The Chronicle of Higher Education*, 1975, 11–1.) reported that "the 1975 high school graduates posted new lows in average scores on the Scholastic Aptitude Test (SAT)." While the yearly decline was the greatest in 1975 (10 points on the verbal part and 8 points on the math part), it is only one step in a 13 year pattern of decline. Many attempts have been made to explain these declines. To date the phenomenon seems real and not just an artificially based change. Similar results have been

APPLYING MATHEMATICS WITHOUT A LICENSE

MELVIN HENRIKSEN[1]

A recent educational experience made me realize the extent to which the mathematical community has become fragmented and how this has served to inhibit communication both with others and ourselves. What used to be called a department of mathematics has split into several parts by adding departments of computer science and statistics for openers, not to mention separate departments of applied mathematics in some places, and various specialized organizations of mathematicians plying their trade in colleges of agriculture, commerce, or engineering under odd sounding names. Undergraduate colleges too small to afford so many different departments make up for this lack by offering options within a major in mathematics which reflect these differences. Since most methods used to finance departments or programs depend heavily on enrollments, walls are erected which serve to make exploration of more than one brand of mathematics by the student difficult; even as an undergraduate.

If the prime educational goal is to train people to push back the frontiers of knowledge in these areas, perhaps a high degree of specialization is needed. But I wonder if even such a noble goal justifies forcing college sophomores to choose a specialty so early that we end up with Ph.D's in computer science who are shaky about the definition of a limit of a function, pure mathematicians who have never been exposed to statistics, and statisticians who have never been exposed to either abstract algebra or topology. Not only do we create a yawning gap between "pure" and "applied" mathematics, but we also maintain sharp differences between various kinds of the latter which demand, seemingly, the undivided attention of the student beginning at an early point in his college studies. It is as if one needs a license to apply mathematics obtainable only after many years of a highly specialized apprenticeship.

Harvey Mudd College is an undergraduate institution with an enrollment of about 450, and (with a small number of exceptions) all of our students major in chemistry, engineering, mathematics or physics. Despite this small size, our program in mathematics has succumbed to many of the trends towards early specialization mentioned above. Through our Mathematics Clinic, however, we have managed to get some of our majors (and ourselves) to apply mathematics to a variety of problems obtained from industrial and governmental organizations. These are not classroom exercises in applicable mathematics. The organization involved pays Harvey Mudd College to assign a team consisting of a faculty member and a group of students to work on an open-ended problem of real concern to them and produce a report on their progress after some pre-assigned time (usually one or two semesters). Students work on clinic projects while taking other courses. The Mathematics Clinic is run in cooperation with the Claremont Graduate School, and participants may come from any of the Claremont Colleges[2].

The Mathematics Clinic began in 1973 and is an offshoot of a corresponding operation started much earlier by our engineering department. The prerequisites for participation in a particular project vary with the problems involved, and different students may contribute different skills. The projects taken on so far include mathematical modeling of air pollution and of an industrial problem involving light scattering, developing a kinematic handbook for missile design, and the mathematical modeling of the history of a compacting sedimentary basin. The mathematical techniques used include

---

[1] This article was prepared while the author was on leave at the University of Manitoba in Winnipeg, Manitoba, Canada.

[2] The Claremont Graduate School, Claremont Men's College, Harvey Mudd College, Pitzer College, Pomona College, and Scripps College form a group of independent colleges located in Claremont, California which share certain facilities in common and cooperate by mutual consent in some ways. Each has is own curriculum and degree requirements, but students are permitted to take some of their courses at a Claremont College other than their own.

computer simulation, ordinary and partial differential equations, the finite element method, probability, statistics, and, of course, linear algebra. The students participating must be able to read background material relevant to a particular project. Hence there is a limit to the sophistication of the techniques that are used.

My colleague George Orland and I supervised a project on juror utilization for the Superior Court of the County of Los Angeles in the spring of 1975. Our goal was to produce a method for reducing the number of people who have to be called to jury duty without slowing down the flow of court cases coming to trial or initiating legislative changes. Both of us are "pure" mathematicians whose research interests are in functional analysis, general topology and the structure of rings. The students[3] ranged from a freshman studying the calculus, through some seniors with knowledge of probability and statistics, and a graduate student aiming for a career in junior college teaching. None of us were certified as applied mathematicians or had any background in law, but we thought we could improve the current jury selection system which kept large numbers of people in attendance each day who never got into a courtroom. Since we had to come up with some answers within a semester, we decided to concentrate our efforts in one courthouse, and chose the one in Pomona, California because of its geographic proximity and the reasonably small volume of cases handled there.

Jumping ahead in the story, we did succeed in devising a rather crude method which, if followed, would reduce the number of jurors needed in the Pomona Courthouse by at least 37% and save the county at least $73,000 per year. Moreover, with more effort, the method could be improved and with modification applied to other courthouses in Los Angeles County. George Orland and I have reported on this in full in a legal journal **[1]**, including our difficulties in getting our system implemented, which continue until now. Only a little of **[1]** is reported below.

Originally we thought we could readily devise a mathematical model, test its effectiveness with the aid of data in court records, modify it accordingly, and refine it by a process of successive approximation to reality. Such data turned out not to be readily available. For example, while we could find out how many people had been sent into a courtroom (usually in groups of 30 to 40) to be examined by attorneys as possible jurors for a particular trial, no record was kept of how many of them were actually questioned before a jury of 12 (sometimes with 1 or 2 alternates) was completed. So we had to gather our own data, and to talk with the jury commissioner's office, attorneys, judges and court clerks to obtain data. To win cooperation we had to explain our purpose, and then listen to an unsolicited discourse on why mathematics could be of no help. One by one, they told us essentially the same story concerning the great variability involved in the jury selection process and emphasized their point with the fact that over 700 prospective jurors had to be examined to get a jury for the murder trial in which Charles Manson was the defendant. Each of them seemed convinced that we were seeking some universal number of jurors that would provide efficiently for the needs of the courthouse every day it is in session. Only by listening to their discourse, and convincing them we were not quite so naive as they had thought, could we get them (sometimes reluctantly) to answer our questions. The overwhelming majority of those whom we questioned held college degrees; yet none of them seemed to realize that mathematical techniques could deal with uncertainty.

The mathematical techniques needed were not at all sophisticated. Computer programming was a vital tool for digesting the data we had gathered. Some elementary statistics was needed to determine the number of jurors that had to be examined to complete a jury to within a pre-assigned confidence interval. Usually, 25 is a good approximation. (Clearly "sensational" trials have to be handled separately. These are known well in advance, are few in number, and do not occur at all in outlying courthouses like the one in Pomona). Since a new jury trial cannot start without an available courtroom, to get an upper bound for the number of jurors needed on the next day, it is enough to multiply the number of courtrooms that have a chance of becoming vacant by 25.

---

[3] John Lavrakas of the Claremont Graduate School, David Abrahamson and Joseph Coquillard of Harvey Mudd College, Deborah Taper of Pitzer College, John Irvine and Greg Johnson of Pomona College.

The interested reader can examine [1] for possible refinements of this scheme. The bulk of the time spent was in interviewing people and gathering data; as opposed to learning about recent developments in applied mathematics. Even this crude approach can save a lot of the taxpayers money and reduce considerably the time wasted by citizens called for jury duty.

As of this writing, we have not been encouraged to refine our techniques, nor are they being used to their full potential, despite accurate publicity on our project by the *Los Angeles Times* on August 21, 1975. I can only conjecture as to why, but two reasons seem very plausible. The budget of the Superior Court of Los Angeles County was cut substantially for 1975–76, and there is an initial cost of any administrative reform which could not be recovered later because of the nature of government financing. (The standard reward for efficiency in public institutions is usually more work and a budget cut). Also, while the mathematics involved seems quite elementary to us, it is less than comprehensible to those who bear the responsibility for changing the system. This latter problem, pointed out so well by C. P. Snow in [2] seems no closer to a solution now than it was then.

What about the rewards to the mathematicians working on problems of this kind? We never considered sending [1] to a mathematical (or statistical or computer science) journal since the tools used are too elementary to add to the technical knowledge of their readers. Surely no young man trying to get tenure today in any academic department in the mathematical sciences at a university can afford to work on a project whose results cannot be published in a mathematical journal. Yet we did apply mathematics to a problem of social importance. Some years ago, while I was on the faculty of a midwestern state university I had a colleague who had co-authored quite a few papers with social scientists. He had applied the elements of probability, statistics, or linear algebra to problems considered important by a variety of psychologists and sociologists. To them, he was the most valuable member of the department of mathematics. His mathematical colleagues gave him little or no credit for these efforts since they considered the mathematical content of the papers trivial, and he left after a while. It is the sophistication of the tool used about which we seem to care, as opposed to the problem on which it is being used.

The renaissance man is no more likely to return to the mathematical scene than the passenger pigeon is to reappear in the skies of North America. Mathematics has grown too much for any individual to keep up with the current literature in all but a small part of it, or to be able to teach at the graduate level in all of the areas we now call the mathematical sciences. But in our zeal to get research done and to train young people to succeed us, we are in danger of becoming *Fachidioten* capable of communicating only with a small group of mathematicians with the same research interests. Indeed, when I find that I cannot read many research articles in fields close to my own because the author won't take the trouble to explain the notation used or give a reference to where it is defined, I wonder if we are not in danger of joining the passenger pigeon.

Without a space race or ever expanding college enrollments to provide a justification for our activities, we must seek other ways to provide employment for young mathematicians than learning how to reproduce themselves. Perhaps teaching them and ourselves how to apply mathematics without the total commitment involved in getting a license can help.

Much more thorough descriptions of the Mathematics Clinic are given by Jerome Spanier in [3] and [4].

References

1. Melvin Henriksen and George Orland, On the juror utilization problem, The Jurimetrics Journal, 16 (1976) 318–333.

2. C. P. Snow, Two Cultures and The Scientific Revolution, Cambridge University Press, New York, 1962.

3. J. Spanier, The Mathematics Clinic: an innovative approach to realism within an academic environment, this MONTHLY, 83 (1976) 771–775.

4. ———, The Claremont Mathematics Clinic, Siam Review, 19 (1977) 536–549.

DEPARTMENT OF MATHEMATICS, HARVEY MUDD COLLEGE, CLAREMONT, CA 91711.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

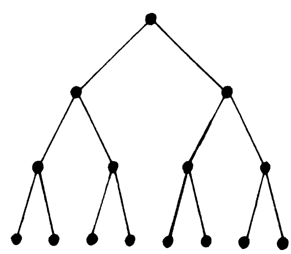
ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before January 31, 1978.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

E 2671.* *Proposed by Ibrahim Cahit, Nicosia, Cyprus*

Let $T = (V, E)$ be a k-level complete binary tree (see figure for $k = 4$) with vertex set V and the



edge set E. Thus $|V| = 2^k - 1$ and we put $N = \{1, 2, 3, \ldots, 2^k - 1\}$. For every bijection $f\colon V \to N$ define

$$W(f) = \sum_{\{i,j\} \in E} |f(i) - f(j)|.$$

Prove or disprove $\min_f W(f) = (k-1)2^{k-1}$ $(k \geqq 2)$.

E 2672. *Proposed by Marianne Gardner, North Carolina State University*

Each of the $\binom{m}{2}$ edges of the complete graph K_m is assigned a direction and each vertex is assigned one of n colors in such a way that there is no directed path of length k $(k < m)$ whose vertices are all of the same color. How large can m be in terms of n and k? (This problem is similar to E 2562 [1977, 218].)

E 2673. *Proposed by Haim Rose, Kiriat Schmonah, Israel*

Let $p = 6n + 1$ be a prime number, n a positive integer. An *n-residue mod p* is an integer a such that $0 < a < p$ and $a \equiv b^n \bmod p$ for some integer b. Prove that the product of all n-residues mod p which are less than $p/2$ is congruent to $-1 \bmod p$.

E 2674.* *Proposed by G. Tsintsifas, Thessaloniki, Greece*

Let $S = \{A_0, A_1, \ldots, A_n\}$ and $S' = \{A'_0, A'_1, \ldots, A'_n\}$ be regular n-simplices such that A'_i lies on the face $\{A_0, \ldots, A_{i-1}, A_{i+1}, \ldots, A_n\}$ of S $(0 \leqq i \leqq n)$. Is it true that the centroids of S and S' coincide?

E 2675. *Proposed by R. P. Boas, Northwestern University*

If n is a positive integer let $f(n)$ be the number of zeros in the decimal representation of n. For which values of $a > 0$ is the following series convergent

$$\sum_{n \geqq 1} \frac{a^{f(n)}}{n^2}?$$

E 2676. *Proposed by Robert Gilmer, Florida State University*

Let R be a ring (not necessarily with identity). We denote by R_n the ring of $n \times n$ matrices over R. Show that the following are equivalent

(i) Every ideal of R_n is of the form I_n where I is an ideal of R,

(ii) $I = IR = RI$ holds for every ideal I of R.

SOLUTIONS OF ELEMENTARY PROBLEMS

Broken-Line Brachistochrone

E 1255 [1957, 109; 1975, 661]. *Proposed by J. P. Ballentine*

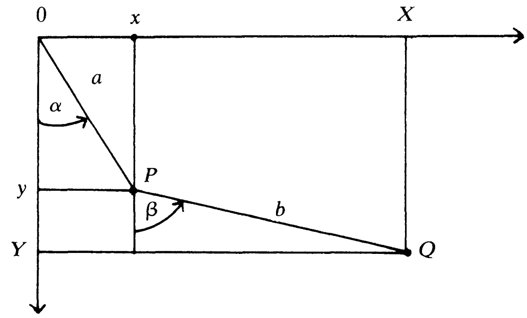
A constant gravitational field operates in the direction of the negative y-axis. A particle starts from point (a, b) with a given initial velocity, travels along a straight line to the point (x, y), thence without loss of velocity (speed) along another straight line to the point (c, d). The point (x, y) is so chosen that the total time is a minimum. Prove or disprove that the particle reaches the point (x, y) when the time is precisely half gone.

Solution by Taw-Pin Lim, Toronto, Ontario. We shall prove the validity of the statement, with notation changed as follows:

Change the orientation of the y-axis;
$O = (0, 0)$ is the initial point;
$P = (x, y)$ is the break-point;
$Q = (X, Y)$ is the terminal point;
v_0 is the initial speed;
v is the speed at the point P; $OP = a$;
v_1 is the speed at the point Q; $PQ = b$;
α is the angle between the positive y-axis and $\overrightarrow{OP}$; t_1 is the transit time for the segment OP;
β is the angle between the positive y-axis and $\overrightarrow{PQ}$; t_2 is the transit time for the segment PQ;
$T = t_1 + t_2$,
g = the acceleration due to the gravitational field.
We have

(1) $$v = v_0 + gt_1 \cos \alpha, \qquad v_1 = v + gt_2 \cos \beta;$$

(2) $$v^2 = v_0^2 + 2gy, \qquad v_1^2 = v_0^2 + 2gY.$$



Considering x, y as independent variables we obtain

$$\partial a/\partial x = x/a = \sin\alpha, \qquad \partial b/\partial x = (x-X)/b = -\sin\beta;$$

$$\partial a/\partial y = y/a = \cos\alpha, \qquad \partial b/\partial y = (y-Y)/b = -\cos\beta.$$

Since $t_1 = 2a/(v_0+v)$ and $t_2 = 2b/(v+v_1)$ we find that

$$\frac{\partial t_1}{\partial x} = \frac{t_1}{a}\sin\alpha, \qquad \frac{\partial t_1}{\partial y} = \frac{t_1}{a}\cos\alpha - \frac{t_1^2 g}{2av};$$

$$\frac{\partial t_2}{\partial x} = -\frac{t_2}{b}\sin\beta, \qquad \frac{\partial t_2}{\partial y} = -\frac{t_2}{b}\cos\beta - \frac{t_2^2 g}{2bv}.$$

Since $T = t_1 + t_2$, the equations $\partial T/\partial x = \partial T/\partial y = 0$ give

(3) $$\frac{t_1}{a}\sin\alpha = \frac{t_2}{b}\sin\beta,$$

(4) $$\frac{t_1}{a}\left(\cos\alpha - \frac{t_1 g}{2v}\right) = \frac{t_2}{b}\left(\cos\beta + \frac{t_2 g}{2v}\right).$$

From (3) we obtain

(5) $$\frac{\sin\beta}{\sin\alpha} = \frac{bt_1}{at_2} = \frac{v+v_1}{v_0+v} = z \quad \text{(say)}.$$

Putting $p = \cos\alpha$, $q = \cos\beta$ we obtain from (4) the equation

(6) $$zq\ [2vp^2 - (v-v_0)] = p[2vq^2 + (v_1 - v)].$$

We have

(7) $$q^2 = 1 - \sin^2\beta = 1 - z^2\sin^2\alpha = 1 - z^2 + p^2z^2.$$

By squaring (6) and using (7) to eliminate q^2 we get

(8) $$(1 - z^2 + uz^2)[2uv - (v - v_0)]^2 = u[2vz(u-1) + (v_0 + v)]^2,$$

where $u = p^2$. The coefficients of u^3 in the two members of (8) are equal. Thus (8) is a quadratic equation in u. Since $u = 1$ is a root of (8) we can easily rewrite (8) in the factorized form

(8′) $$(u-1)[4uv(v_1 - v_0z^2) - (z^2-1)(v-v_0)^2] = 0.$$

It is easy to see that, in general, we have $u \neq 1$ and hence we must have

(8″) $$4uv(v_1 - v_0z^2) = (z^2-1)(v-v_0)^2.$$

Since

$$(v_1 - v)^2 - z^2(v - v_0)^2 = (v_1 + v)^2 - 4vv_1 - z^2(v + v_0)^2 + 4vv_0z^2$$
$$= 4v(v_0z^2 - v_1),$$

we obtain from (8″) and (7) that

$$(v_1 - v)^2p^2 = [z^2(v - v_0)^2 + 4v(v_0z^2 - v_1)]u$$
$$= z^2(v - v_0)^2p^2 - (z^2 - 1)(v - v_0)^2 = (v - v_0)^2q^2.$$

Thus

$$\left|\frac{v_1 - v}{q}\right| = \left|\frac{v - v_0}{p}\right|$$

which, in view of (1), gives $t_1 = t_2$.

Also solved by Peter Ungar.

A partial solution, for $v_0 = 0$, was submitted by David Wright.

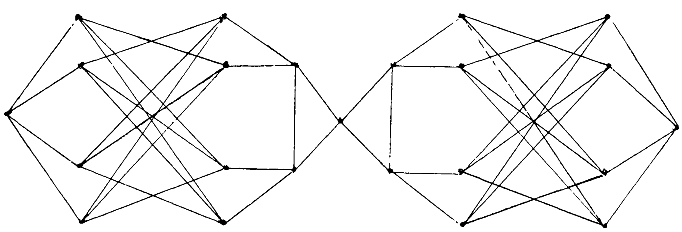
Covering Vertices of Four-valent Graphs

E 2564 [1975, 1009; 1977, 214]. *Submitted by D. J. Kleitman, Massachusetts Institute of Technology*

Can one cover the vertices of any regular graph of degree 4 (every vertex in it has degree four) by disjoint arcs and stars? (A star consists of a vertex and all the arcs containing it; an arc covers both its ends.) [This problem is due to R. L. Graham and is a variation of a problem of Claude Berge.]

Recall of a solution. The published solution (March 1977) does not solve the problem as proposed. The problem was misinterpreted on two accounts: first, a *star* was considered as consisting of one vertex only and four incident edges; second, instead of covering only vertices we insisted that the edges should be covered. We are indebted to Jerry Griggs and Daniel Kleitman for this correction.

Solution by Daniel J. Kleitman. The following is a counterexample. It is easy to verify that no such covering is possible.



Subadditive and Superadditive Numbers

E 2590 [1976, 284]. *Proposed by C. A. Nicol, University of South Carolina*

A natural number $n \geqq 2$ is said to be ϕ-subadditive if $\phi(n) \leqq \phi(k) + \phi(n-k)$ for $1 \leqq k \leqq n-1$ and ϕ-superadditive if $\phi(n) \geqq \phi(k) + \phi(n-k)$ for $1 \leqq k \leqq n-1$ (ϕ denotes Euler's totient function). Show that there exist infinitely many ϕ-subadditive numbers and infinitely many ϕ-superadditive numbers.

Solution by Lorraine L. Foster, California State University at Northridge. Let $p \geqq 5$ be prime. Now $1 < k < p$ implies $\phi(k) + \phi(p-k) \leqq (k-1) + (p-k-1) = p-2 < \phi(p)$. Also, $p-1$ is not prime so that $\phi(p-1) \leqq p-3$ and $\phi(1) + \phi(p-1) \leqq 1 + p - 3 < \phi(p)$. Hence the primes greater than 3 are ϕ-superadditive.

Let $n_r = p_1 p_2 \cdots p_r$ be the product of the first r primes, $r \geqq 2$. Let $1 < m < n_r$ and let $q_1, q_2, \ldots, q_s$ be the prime divisors of m. Then

$$\frac{\phi(m)}{m} = \prod_{i=1}^{s}\left(1-\frac{1}{q_i}\right) \geqq \prod_{i=1}^{s}\left(1-\frac{1}{p_i}\right) > \prod_{i=1}^{r}\left(1-\frac{1}{p_i}\right) = \frac{\phi(n_r)}{n_r}.$$

Also $\phi(1)/1 > \phi(n_r)/n_r$. Hence $1 \leqq k < n_r$ implies that

$$\phi(k) + \phi(n_r - k) > k\phi(n_r)/n_r + (n_r - k)\phi(n_r)/n_r = \phi(n_r).$$

Thus the n_r are ϕ-subadditive.

Also solved by David Anick & Ira Gessel, David Bienenfeld (Israel), Robert Breusch, David Carlson, José Luis de Miguel (Spain), Miltiades Demos, M. G. Greening (Australia), J. D. Jones, Lael Kinch, L. Kuipers (Switzerland) & H. Niederreiter, Jordan Levy, Graham Lord (Canada), Robert Patenaude, Barry Powell, Daniel Rosenblum, I. J. Schoenberg, University of South Alabama Problem Group, G. W. Valk, and the proposer. Partial solutions were given by Michael Ecker, and Thomas Elsner.

Editor's Comment. Powell showed, with minor modifications of the above proof, that $\phi(k)+\phi(n-k)$ may be replaced in the problem by $\Sigma_{i=1}^{m} \phi(s_i)$ where $\Sigma_{i=1}^{m} s_i = n$. It seems difficult to characterize either the superadditive or the subadditive numbers.

Null Sequences and Convergent Series

E 2591 [1976, 284]. *Proposed by Jan Mycielski, University of Colorado*

Prove that for every real sequence $\{a_n\}$ with $\lim a_n = 0$ there exists a sequence $\{b_n\}$ with $b_1 \geqq b_2 \geqq \cdots \geqq 0$ such that $\Sigma\, b_n$ diverges and $\Sigma\, a_n b_n$ converges absolutely.

Solution by Charles E. Blair, University of Illinois, Urbana. Define a sequence of natural numbers $\{m_k\}$ such that

$$\text{(i) } |a_n| < 2^{-k} \text{ if } n \geqq m_k, \qquad \text{(ii) } m_k > 2m_{k-1}.$$

For $m_k < n \leqq m_{k+1}$ define $b_n = 1/(m_{k+1} - m_k)$. By (ii), $m_{k+1} - m_k > m_k - m_{k-1}$, so $b_{k+1} \leqq b_k$ for all k. Since

$$\sum_{m_k+1}^{m_{k+1}} b_n = 1,$$

$\Sigma\, b_n$ diverges. Since

$$\sum_{m_k+1}^{m_{k+1}} |a_n b_n| = \{1/(m_{k+1} - m_k)\} \sum_{m_k+1}^{m_{k+1}} |a_n| \leqq 2^{-k},$$

$\Sigma\, a_n b_n$ converges absolutely.

Also solved by George Akst, K. F. Andersen (Canada), Gary Bates, David Bienenfeld (Israel), Robert Brigham, F. S. Cater, P. G. Chauveheid (Belgium), Michael Dixon, Robert Gilmer, Danny Goldstein, Gustaf Gripenberg (Finland), William Habakkuk, Ellen Hertz, Ole Jørsboe (Denmark), Benjamin Klein, Joel Levy, Jordan Levy, O. P. Lossers (Netherlands), MIT Combinatorics Class, Walter Maxey, Adam Riese, Michael Skalsky, St. Olaf Problems Group, Arthur Solomon & C. Belna & J. Brown, Allen Stenger, University of South Alabama Problem Group, Laird Taylor, J. B. Wilker (England), Abraham Ziv (Israel), and the proposer.

Editorial Note. A closely related assertion of Stieltjes can be found on page 47 of Bromwich, *An Introduction to the Theory of Infinite Series*, Macmillan and Co. Ltd., London 1955.

An Application of Cayley's Theorem

E 2592 [1976, 285]. *Proposed by Melvin Hausner, New York University*

Let G be a finite group of even order $n = 2m$. Let H be the set of all x in G with $x^m = 1$. Prove (a) H is a subgroup of G and (b) either $H = G$ or the index $[G:H]$ is 2.

Solution. Let $\theta\colon G \to \mathrm{Sym}(G)$ be the Cayley (left regular) representation. To prove (a) and (b) it suffices to show that $\theta(H) = \theta(G) \cap \mathrm{Alt}(G)$. Let $x \in G$, let k be the order of x and let $s = 2m/k$. Then $\theta(x)$ is a product of s disjoint k-cycles. Since $sk = 2m$, it follows that $\theta(x)$ is even if and only if s is even, i.e., $\theta(x) \in \mathrm{Alt}(G)$ if and only if $x \in H$.

Such solutions were submitted by Michael Barr (Switzerland), Lee Erlebach, Lorraine Foster, Richard Goldstein, M. G. Greening (Australia), O. P. Lossers (Netherlands), M. I. T. Combinatorics Class (Ted Chinburg & Jerry Griggs), Roberto Mena, and the proposer.

Also solved by the Bennett College Team, D. M. Bloom, Sidney Heller, Mark Hopkins, A. A. Jagers (Netherlands), Janet Locke, L. E. Mattics, Louise Moser, Eric Rosenthal, Sidney Schipper, Scrandis Playtis, Rony Teitler (Belgium), and S. G. Udpikar & M. R. Modak (India).

Editor's Comment. Many solvers use the following arguments. If a Sylow 2-subgroup P of G is not cyclic then $H = G$. Otherwise, by Burnside's theorem, P has a normal complement, say N, and then $H = NP_1$ where P_1 is the subgroup of P of index 2. Hopkins and Schipper note that the same arguments prove the following more general result: Let G be a finite group of order $n = pm = p^{k+1}r$ where $p < q$ for all primes q which divide r. Let $H = \{x \in G \mid x^m = 1\}$. Then H is a subgroup of G and $(G\colon H) = p$ or 1 according to whether a Sylow p-subgroup P of G is cyclic or not.

Tiling by Trominoes

E 2595 [1976, 379]. *Proposed by Sidney Penner, Bronx Community College of the City University of New York*

Consider $(2n+1)^2$ hexagons arranged in a "diamond" pattern, the kth column from the left and also the kth column from the right consisting of k hexagons each, $1 \leqq k \leqq 2n+1$ (Figure 1 illustrates the 5^2 case). Show that if $n \not\equiv 1 \pmod 3$ and the center hexagon is deleted, then the remaining hexagons can be tiled by trominoes as in Figure 2.

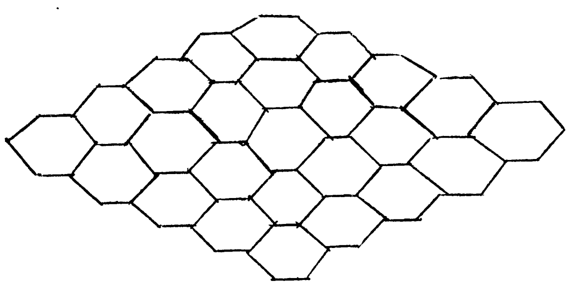


FIG. 1

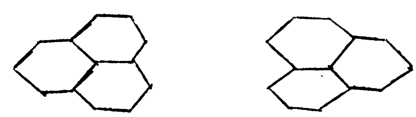


FIG. 2

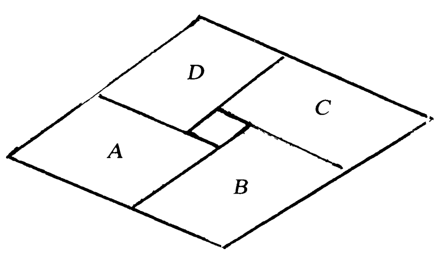


FIG. 3

Solution by Eli Leon Isaacson, New York University. The diamond can be split into four congruent $n\times(n+1)$ "boxes" A, B, C, D as in Figure 3. It suffices to show that A can be tiled by trominoes.

Since $n\not\equiv 1 \pmod 3$, we have $3|n$ or $3|(n+1)$. A 2×3 box can be split in two trominoes. Hence if $3|n$ and $n+1$ is even or $3|(n+1)$ and n is even, then the $n\times(n+1)$ box can be tiled.

If $3|n$ and n is even then we split the box A into two boxes of sizes $n\times(n-2)$ and $n\times 3$. Each of these two boxes can be split into 2×3 boxes. A similar argument applies to the case $3|(n+1)$ and n odd.

Also solved by Robert Brigham, John De Carlo, Thomas Foregger, Michael Goldberg, William Gould, Keith Hodge, Elgin Johnston, Jordan Levy, Steven Locke (Canada), Aaron Meyerowitz (Israel), Zoran Taborin (Yugoslavia), G. W. Valk, and the proposer.

Transformation Induced in a Symmetric Power

E 2597 [1976, 379]. *Proposed by R. W. Farebrother, University of Manchester, England*

Let j, n be integers such that $0\leqq j\leqq n$ and let

$$(1-x)^j(1+x)^{n-j}=\sum_{i=0}^{n} c_{ij}(n)x^i.$$

If $C(n)$ is the matrix $(c_{ij}(n))$ where i, $j=0,1,\ldots,n$ show that

$$C(n)^2=2^n\cdot I,$$

$$\det C(n)=(-2)^{n(n+1)/2}$$

$$\operatorname{tr} C(n)=0 \text{ if } n \text{ is odd}$$

$$=2^{n/2} \text{ if } n \text{ is even.}$$

I. *Solution by Ladnor Geissinger, University of North Carolina.* It is immediate from the definition that

(1) $$(1,x,\ldots,x^n)C=(1+x)^n\cdot\left(1,\frac{1-x}{1+x},\ldots,\left(\frac{1-x}{1+x}\right)^n\right).$$

Multiplying by C and using (1) we get

$$(1,x,\ldots,x^n)C^2=(1+x)^n\left(1+\frac{1-x}{1+x}\right)^n(1,x,\ldots,x^n)$$

$$=2^n(1,x,\ldots,x^n).$$

This gives $C^2=2^n\cdot I$.

Let $x_0,x_1,\ldots,x_n$ be distinct numbers and $y_i=(1-x_i)/(1+x_i)$. We denote by $V(x_0,x_1,\ldots,x_n)$ the Vandermonde matrix whose kth row is 1, $x_k,\ldots,x_k^n$, $0\leqq k\leqq n$. It follows from (1) that

(2) $$V(x_0,x_1,\ldots,x_n)C=(I+D)^nV(y_0,y_1,\ldots,y_n),$$

where $D=\operatorname{diag}(x_0,x_1,\ldots,x_n)$. We have

$$\det V(y_0,y_1,\ldots,y_n)=\prod_{i>j}(y_i-y_j)$$

$$=(-2)^{n(n+1)/2}\left(\prod_{i=0}^{n}(1+x_i)\right)^{-n}\prod_{i>j}(x_i-x_j).$$

By taking determinants of both members of (2) we obtain the required formula for $\det C$.

Assume that $n = 2k$ is even and that the x_i were chosen so that $x_0, y_0, x_1, y_1, \ldots, x_{k-1}, y_{k-1}, x_k$ are distinct and $y_k = x_k = \sqrt{2} - 1$. Then the vectors

$$\mathbf{v}_i = (1, x_i, \ldots, x_i^n), \qquad 0 \leqq i \leqq k$$

$$\mathbf{w}_i = (1, y_i, \ldots, y_i^n), \qquad 0 \leqq i \leqq k-1$$

form a basis of $\mathbf{R}^{n+1}$ and (1) gives

$$\mathbf{v}_i C = (1 + x_i)^n \mathbf{w}_i, \qquad 0 \leqq i \leqq k-1$$

$$\mathbf{w}_i C = (1 + y_i)^n \mathbf{v}_i, \qquad 0 \leqq i \leqq k-1$$

$$\mathbf{v}_k C = (1 + x_k)^n \mathbf{v}_k = (\sqrt{2})^n \mathbf{v}_k.$$

Therefore $\operatorname{tr} C = (\sqrt{2})^n$.

If n is odd a similar argument gives $\operatorname{tr} C = 0$. Alternately, $C^2 = 2^n \cdot I$ implies that the eigenvalues of C are $\lambda_i = \varepsilon_i(\sqrt{2})^n$ where $\varepsilon_i = \pm 1$, $0 \leqq i \leqq n$. Since

$$\operatorname{tr} C = (\varepsilon_0 + \varepsilon_1 + \cdots + \varepsilon_n)(\sqrt{2})^n$$

has to be rational, if n is odd, we must have $\varepsilon_0 + \varepsilon_1 + \cdots + \varepsilon_n = 0$.

II. *Solution by Clark Givens, Michigan Technological University (revised by the editor)*. Let $T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ considered as a linear transformation of the 2-dimensional complex space $V = \mathbf{C}^2$. Let T_n be the induced transformation acting on the nth symmetric power $S^n(V^*)$ of the dual space V^*.

If x and y are coordinate functions of V, with respect to the standard basis, then $x^j y^{n-j}$, $0 \leqq j \leqq n$, is a basis of $S^n(V^*)$ and

$$T_n(x^j y^{n-j}) = (ax + cy)^j (bx + dy)^{n-j}.$$

The entries of the matrix of T_n, say $A = (\alpha_{ij})$, are given by

$$(ax + cy)^j (bx + dy)^{n-j} = \sum_{i=0}^{n} \alpha_{ij} x^i y^{n-i}, \qquad 0 \leqq j \leqq n.$$

Let λ_0, λ_1 be the eigenvalues of T. Then it is well known that the eigenvalues of T_n are

$$\lambda_{i_1} \lambda_{i_2} \ldots \lambda_{i_n} \qquad (0 \leqq i_1 \leqq i_2 \leqq \cdots \leqq i_n \leqq 1),$$

i.e., $\lambda_0^k \lambda_1^{n-k}$, $0 \leqq k \leqq n$. Thus we have

$$\operatorname{tr} T_n = \begin{cases} \dfrac{\lambda_1^{n+1} - \lambda_0^{n+1}}{\lambda_1 - \lambda_0} & \text{if } \lambda_1 \neq \lambda_0 \\[2ex] (n+1)\lambda_0^n & \text{if } \lambda_1 = \lambda_0 \end{cases}$$

$$\det T_n = (\lambda_0 \lambda_1)^{n(n+1)/2} = (ad - bc)^{n(n+1)/2}.$$

If $T = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$ we have $T_n = C(n)$, $\lambda_0 = \sqrt{2}$, $\lambda_1 = -\sqrt{2}$ and the above formulas give the required expressions for $\det C(n)$ and $\operatorname{tr} C(n)$. Since $T^2 = 2 \cdot I$ we also have

$$(T_n)^2 = (T^2)_n = (2 \cdot I)_n = 2^n \cdot (I)_n = 2^n \cdot I.$$

Also solved by M. T. Bird, Robert Breusch, Robert Brigham, Paul Bruckman, Leonard Carlitz, F. Gerrish (England), M. G. Greening (Australia), Eli Isaacson, Robert Kleinhenz & Elgin Johnston, I. I. Kolodner, José Luis de Miguel (Spain), L. E. Mattics, David Rubin, Yahya Said (Iraq), and the proposer.

Editor's Comment. The advanced problem 6010 [1976, 666] dealt also with the coefficients $c_{ij}(n) = c^{(i)}_{n-j,j}$. The proposer of that problem computed the sums

$$\sum_{k=0}^{t} c^{(k)}_{m,n}\, c^{(k)}_{t-q,q} = \sum_{k=0}^{t} c_{kn}(t)c_{kq}(t).$$

This amounts to computing the entries of $C(t)' \cdot C(t)$ where $'$ denotes the transpose.

Dense Rational Set with Irrational Distances

E 2598 [1976, 379]. *Proposed by Erwin Just, Bronx Community College of CUNY*

Does there exist a set of rational points which is dense in the plane such that the distance between each pair of points in the set is irrational?

Solution by Peter L. Montgomery, Huntsville, Alabama. The answer is positive. Let (x_i, y_i), $i = 0, 1, 2, \ldots$ be a sequence of points which is dense in the plane. For each i we can choose odd integers a_i and b_i such that

$$|x_i - 2^{-i}a_i| \leqq 2^{-i}, \qquad |y_i - 2^{-i}b_i| \leqq 2^{-i}.$$

Then the sequence of rational points $P_i = (2^{-i}a_i, 2^{-i}b_i)$, $i = 0, 1, 2, \ldots$ is also dense in the plane. If $i < j$ then

$$4^j\,\overline{P_iP_j}^2 = (a_i^2 + b_i^2)4^{j-i} - (a_ia_j + b_ib_j)2^{j-i+1} + (a_j^2 + b_j^2).$$

Since the right-hand side is congruent to $2 \bmod 4$, $\overline{P_iP_j}$ is irrational.

Also solved by Irl Bivens, William Gorman, Elgin Johnston, Peter Lindstrom, L. E. Mattics, L. F. Meyers, Adam Riese, J. G. Wendel, and the proposer.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Hill Center, Rutgers University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate signed sheets and should be mailed before January 31, 1978.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

6168. *Proposed by Edmond Dale Dixon, Tennessee Technological University*

The following theorem appears in several recent books, e.g. C. G. Cullen, *Matrices and Linear Transformations*; J. D. Gilbert, *Elements of Linear Algebra*; and Ben Noble, *Applied Linear Algebra.* "Let A be a diagonalizable matrix with eigenvalues $\lambda_1, \lambda_2, \ldots$, such that $|\lambda_1| > |\lambda_2| \geqq \cdots$, and let X be any vector not in the subspace spanned by the eigenvectors associated with $\lambda_2, \lambda_3, \ldots$. Let E_i be the vector with 1 in the ith position and zeros elsewhere. Then $E_i \cdot A^{n+1}X / E_i \cdot A^nX \to \lambda_1$, for each i, where the denominators are nonzero."

Find a counterexample.

6169. *Proposed by Joseph Rotman, University of Illinois*

Prove that the category of all Lie algebras over a field K has no injective objects other than 0.

6170. *Proposed by Paul W. Haggard, East Carolina University*

Let D be an integral domain with characteristic the prime p and let x and y be indeterminates. In $D[x, y]$, consider expansions of $(x + y)^n$ for nonnegative integers n.

(a) If p_1 is an odd prime, prove that the expansion of $(x + y)^{p_1}$ has an even number, N, of terms.

(b) When and how can n be obtained such that the expansion of $(x + y)^n$ will have a given number, N, of terms?

6171. *Proposed by R. W. K. Odoni and J. B. Wilker, the University of Exeter, England*

Let F be a field and let n and d be positive integers, each $\geqq 2$. Let σ be any permutation of $\{1, 2, \ldots, n\}$ and let σ_0 be the n-cycle $j \to j+1 \pmod n$. Prove that σ is a power of σ_0 if and only if for every sequence of n $d \times d$ matrices over F, trace $\Pi_{j=1}^n M_j = \operatorname{trace} \Pi_{j=1}^n M_{\sigma(j)}$.

6172*. *Proposed by Doug Hensley, Institute for Advanced Study, Princeton, New Jersey*

Give an example, if possible, of two planar lattices of unit determinant that do not possess a common bounded measurable fundamental domain. Do any two distinct lattices possess a common fundamental domain?

6173. *Proposed by Otomar Hájek, Case Western Reserve University*

For C^2 functions $f \neq 0$ vanishing at 0 and π, consider the functional $\inf_{(0,\pi)} f''/f$ (ignore undefined values). Show that its maximum -1 is attained only by $\sin x$ and its multiples.

SOLUTIONS OF ADVANCED PROBLEMS

A Strong Fermat's Last Theorem

6066 [1976, 62]. *Proposed by the late C. W. Anderson*

For $n = 3$ and $x \in (0,1)$ rational, show that $f_n(x) = (1 - x^n)^{1/n}$ is algebraic of degree n.

Solution by F. J. Flanigan, San Diego State University. We give a solution for many $n \geqq 3$. The problem is related to solvability of the Fermat equation $u^p + v^p = 1$ in rational u, v. To see this, note that $f_n(x)$ is algebraic of degree $\leqq n$, since it is a zero of the polynomial

$$F(n, x; U) = U^n - (1 - x^n) \in \mathbb{Q}[U].$$

Moreover, it is standard that the degree of $f_n(x)$ equals n if and only if $F(n, x; U)$ is irreducible in $\mathbb{Q}[U]$.

By the lemma below, $F(n, x; U)$ is reducible in $\mathbb{Q}[U]$ if and only if $n = pt$ for some prime p and $1 - x^n = y^p$ for some rational y; that is $(x^t)^p + y^p = 1$. Thus we have the following general (but not most general) statement.

THEOREM. *Let $n \geqq 1$. Then for each prime divisor p of n the Fermat equation $u^p + v^p = 1$ has only trivial solutions in $\mathbb{Q}$ if and only if, for all $x \in (0,1)$ rational, the degree of $f_n(x)$ is exactly n.*

It follows that if $n = 3$ or, more generally, if n is any product of odd regular (in the sense of Kummer) primes and/or odd primes $p \leqq 4001$ (among others for which Fermat's Last "Theorem" has been proved), then $f_n(x)$ has degree n.

Of course, even if the Fermat conjecture fails for certain prime divisors p of n (the only known such p is 2), certain $f_n(x)$ may still have degree n.

LEMMA. *Let $n \geqq 1$ and w be a positive rational number, then $U^n - w$ is reducible in $\mathbb{Q}[U] \Leftrightarrow w = z^p$ for a positive rational z and prime divisor p of n.*

Proof. ($\Leftarrow$) $U^n - w = (U^t)^p - z^p$ has the factor $U^t - z$ in $\mathbb{Q}[U]$.

($\Rightarrow$) Let θ be the unique positive real nth root of w, $n \geqq 2$. Then $U^n - w = (U - \theta)(U - \theta\zeta) \cdots (U - \theta\zeta^{n-1})$ in $\mathbb{C}[U]$, where $\zeta = \exp(2\pi i/n)$. Suppose $U^n - w = f(U)g(U)$ nontrivially in $\mathbb{Q}[U]$. Then one has rational constant terms

$$f(0) = (-1)^r\theta^r \cdot \rho, \qquad g(0) = (-1)^s\theta^s \cdot \sigma,$$

where ρ and σ are certain products of the ζ^i, and $f(U)$, $g(U)$ have degrees r, s respectively. Moreover, r, $s < n$. One readily checks that all this forces θ^r to be rational and $\rho = \pm 1$.

Now let $k<n$ be the least integer $\geqq 1$ such that θ^k is rational. A standard argument on remainders shows that k divides n properly. Thus $n = kmp$ for some $m \geqq 1$ and some prime p. Putting $z = \theta^{km}$ yields the assertion.

Also solved by A. A. Jagers (Netherlands), Steve Johnson, Ivan Korec (Czechoslovakia), Roger Lyndon, L. E. Mattics, Dmetri Nakassis, and Valerian Nita.

Negative Values of $\Gamma(z)$

6067 [1976, 62]. *Proposed by Ron Evans, University of Wisconsin, Madison*

Prove that for each real σ, there exist infinitely many $t>0$ for which $\Gamma(\sigma+it)<0$, where Γ denotes the Gamma function.

Solution by A. McD. Mercer, University of Guelph, Ontario, Canada. The asymptotic estimate

$$\log\Gamma(z+a)=(z+a-\tfrac{1}{2})\log z - z + \tfrac{1}{2}\log(2\pi)+o(1) \qquad (|z|\to\infty)$$

provided $|\arg z|\leqq \pi-\delta$, $|\arg(z+a)|\leqq \pi-\delta$ is well known. In this put $a=\sigma$, $z=it$ and equate imaginary parts and we get

$$\arg\Gamma(\sigma+it)=(\sigma-\tfrac{1}{2})\pi/2+t\log t - t + o(1) \qquad (t\to+\infty).$$

Hence any fixed $\sigma>0$ we see that $\arg\Gamma(\sigma+it)$ is equal to an odd multiple of π infinitely often. Hence $\Gamma(\sigma+it)<0$ is satisfied infinitely often on any line $\sigma=\text{constant}>0$, $t>0$.

Also solved by Nathaniel Grossman, Mark Sharlow, and the proposer.

On the Jacobson Radicals

6068 [1976, 62]. *Proposed by Seth Warner, Duke University*

Let A be an algebra over a commutative ring K, and let A_+ be the K-algebra $K\times A$, where addition and scalar multiplication are defined componentwise and multiplication by $(x,a)(y,b)=(xy, x\cdot b+y\cdot a+ab)$. Let N and R be respectively the (Jacobson) radicals of K and A. It is standard that if $N=(0)$, $N\times R$ is the radical of A_+. What are the necessary and sufficient conditions for $N\times R$ to be the radical of A_+?

Solution by William G. Leavitt, University of Nebraska-Lincoln. If J is the Jacobson radical of A_+ then $J=N\times R$ if and only if $NA\subseteq R$.

Proof. Since R is an algebra ideal of A, $0\times R\cong R$ is a radical ideal of A_+. Then since $N\cong(N\times R)/(0\times R)$ is radical, $N\times R$ is also radical. But $NA\subseteq R$ implies $N\times R$ is an ideal of A_+ so $N\times R\subseteq J$. Also the Jacobson radical is hereditary so $J\cap(0\times A)=0\times R$. On the other hand, if $(k,a)\in J$ then $(k,a)\to k\in N$ under the mapping $A_+\to K\cong A_+/(0\times A)$. But $(k,b)\in J$ for any $b\in R$ so $(0,a-b)\in J\cap(0\times A)=0\times R$. Thus $a\in R$, that is $J\subseteq N\times R$. Note that the proof goes for any hereditary radical and in some cases (such as for the nilradical) the condition $NA\subseteq R$ would be automatic.

Also solved by Mrs. D. A. Alamelu (India), John Bryant & Robert Gilmer, Francis Flanigan, Robert Gilmer, Keith Nicholson (Canada), and the proposer.

Zero Divisors and Units in a Group Ring

6069 [1976, 62]. *Proposed by A. R. Charnow, California State University at Hayward*

Let R be an integral domain, G a torsion free group, and $R[G]$ the group ring of G over R. Let $x=r_1g_1+r_2g_2$, $r_i\in R$, $r_i\neq 0$, $g_i\in G$, $g_1\neq g_2$. Prove that x is neither a zero divisor nor a unit in $R[G]$.

Solution by Robert Gilmer, Florida State University. Since

$$r_1g_1 + r_2g_2 = g_1(r_1 + r_2g_1^{-1}g_2) = (r_1 + r_2g_2g_1^{-1})g_1,$$

where g_1 is a unit of $R[G]$, it suffices to prove that if r and s are nonzero elements of R and if g is a nonidentity element of G, then $r + sg$ is neither a unit nor a zero divisor of $R[G]$. Thus, take a nonzero element $f = r_1g_1 + \cdots + r_kg_k$ of $R[G]$, where each r_i is nonzero. Then

$$(r + sg)f = \sum_{i=1}^{k} rr_ig_i + \sum_{i=1}^{k} sr_igg_i,$$

and each of the sets $A = \{g_i\}_{i=1}^{k}$ and $gA = \{gg_i\}_{i=1}^{k}$ has cardinality k. It follows that if $A \neq gA$, then the product $(r + sg)f$ contains at least two nonzero terms, and is therefore neither 0 nor 1. On the other hand, if $A = gA$, then $g_1 = gg_{i_1} = ggg_{i_2} = gggg_{i_3} = \cdots$ for some $i_1, i_2, \ldots$ in $\{1,2,\ldots,k\}$. Consequently, there exist integers m, n with $m < n$ such that $i_m = i_n$ so that $g^mg_{i_m} = g^ng_{i_n}$ and $g^{n-m} = 1$, contrary to the hypothesis that G is torsion free. Therefore $r + sg$ is not a unit or a left zero divisor, and a similar proof shows that $r + sg$ is not a right zero divisor.

Also solved by A. J. Douglas (England), Francis Flanigan, Howard Hiller, A. A. Jagers (Netherlands), Ernst Kani (Germany), Roger Lyndon, Harald Niederreiter, Thomas Sadler, Martin Schechter, and the proposer.

Editor's comment. Flanigan in his solution replaces the condition "G is torsion free" and proves the following: For $y \in R[G]$, define length $(y) = 0$ if $y = 0$; and, if $y \neq 0$ has the unique representation $y = s_1k_1 + \cdots + s_mk_m$ with $m \geqq 1$, the k_i pairwise distinct in G, and all s_i nonzero in R, define length $(y) = m$.

THEOREM. *Let G be any group and g, $h \in G$. Then gh^{-1} is not a torsion element of $G \Leftrightarrow$ for every integral domain R, every $x = ag + bh \in R[G]$ with a, b nonzero in R, and every $y \in R[G]$ with length$(y) = m \geqq 1$, we have length$(xy) \geqq 2$ and also length$(yx) \geqq 2$. In particular, if gh^{-1} is not a torsion element in G, then $x = ag + bh$, with a, b nonzero in R, is neither a unit nor a zero divisor in $R[G]$.*

Counting b for which $\phi(n)/n = a/b$

6070 [1976, 62]. *Proposed by Paul Erdös and the late C. W. Anderson*

Where $\phi(n)$ is Euler's totient function, let $\Phi(n) = \phi(n)/n$, $\Phi\colon N \to (0,1]$ densely. For given a demonstrate that there are only a finite number of b (coprime with a) such that $\Phi(n) = a/b$ has solutions.

Solution by Robert E. Shafer, Berkeley, California. For any n, $n > 1$, we have

$$n = p_1^{a_1}p_2^{a_2}p_3^{a_3}\cdots p_k^{a_k}, \qquad a_i \geqq 1.$$

This gives

$$\Phi(n) = \frac{\phi(n)}{n} = \frac{(p_1 - 1)(p_2 - 1)(p_3 - 1)\cdots(p_k - 1)}{p_1 \quad p_2 \quad p_3 \quad \cdots \quad p_k} = \frac{a}{b}.$$

Each numerator term $p_i - 1$ with the exception of $p = 2$ has a divisor 2. Therefore if a is divisible by 2^{k-2}, then n has at most k distinct prime divisors. Then

$$\Phi(n) \geqq \prod_{2 \leqq p_i \leqq q} \left(1 - \frac{1}{p_i}\right), \qquad q \text{ being the } k\text{th prime}.$$

Thus

$$b \leqq a \Big/ \prod_{2 \leqq p_i \leqq q} \left(1 - \frac{1}{p_i}\right),$$

completing the proof.

The problem was conjectured as long ago as 1972 in conversations between Professor Anderson and myself.

Also solved by Paul Bateman, David Bienenfeld (Israel), Neal Felsinger, O. P. Lossers (Netherlands), L. Kuipers (Switzerland), L. E. Mattics, and the proposers.

Editor's comments. (1) The proposers conjecture that if $f(a)$ is the number of b for which $\Phi(n) = a/b$ has solutions, then $f(a) = o(a^{\varepsilon})$ for all $\varepsilon > 0$. They also raise a similar question for the function $\sigma(n)/n$ where $\sigma(n)$ is the sum of the divisors of n.

(2) Professor Claude Anderson was a frequent contributor to this section with many interesting and significant problems. This was the last prior to his untimely death.

Analytic Mappings of the Unit Disk on a Convex Domain

6071 [1976, 62]. *Proposed by J. G. Milcetich, Federal City College, Washington, D. C.*

The set of analytic functions defined on the unit disk, U, with the topology of uniform convergence on compact subsets of U forms a locally convex, linear topological space. In such a space, $\overline{\text{co}}\, B$ denotes the closed convex hull of a subset B. Let K denote the set of analytic functions $f(z) = z + \sum_{n=2}^{\infty} a_n z^n$, which map U onto a convex domain. Show that for $k \geqq 2$, $z + a_k z^k \in \overline{\text{co}}\, K$ if and only if $|a_k| \leqq \frac{1}{2}$.

Solution by David D. Brannan, University of London, England. Brickman, MacGregor and Wilken [*Convex hulls of some classical families of functions*, Trans. Amer. Math. Soc. 156 (1971), 91–107] have shown that $f(z) = z + a_2 z^2 + \cdots$ belongs to $\overline{\text{co}}\, K$, the closed convex hull of the class K of normalized analytic functions convex in the unit disk U, if and only if $f(z)$ has the representation $f(z) = \int_{|x|=1} z d\mu(x)/(1 - xz)$ where $\mu(e^{it}) \in M[0, 2\pi]$, the class of increasing functions on $[0, 2\pi]$ with $\mu(0) = 0$, $\mu(2\pi) = 1$. There is also a result of Herglotz (see Rudin, *Real and complex Analysis*, McGraw-Hill (1966)) that the function $p(z) = 1 + p_1 z + \cdots$, analytic in $|z| < 1$ belongs to the class P of normalized functions of positive real part in $|z| < 1$ if and only if $p(z) = \int_{|x|=1} (1 + xz) d\mu(x)/(1 - xz)$ for some $\mu(e^{it}) \in M[0, 2\pi]$.

It follows that $f(z)/z \in \overline{\text{co}}\, K$ if and only if $p(z) = 2z^{-1}f(z) - 1 \in P$. The problem seeks necessary and sufficient conditions on a_k that $f_k(z) = z + a_k z^k \in \overline{\text{co}}\, K$; from the above remarks this happens if and only if $1 + 2a_k z^{k-1} \in P$, and so clearly if and only if $|a_k| \leqq \frac{1}{2}$.

We refer to Tsuji, *Potential Theory in Modern Function Theory*, Tokyo (1959) for more information on the coefficient regions for P, which in turn can be used for additional statements about $\overline{\text{co}}\, K$.

Also solved by F. W. Hartmann, Glenn Schober, and the proposer.

Note. The proposer observes that a consequence of the problem is that $z + a_k z^k$ is in the closed convex hull of normalized starlike functions iff $|a_k| \leqq k/2$.

MISCELLANEA

3. "The study of mathematics is so wrongly divorced from the practice of speaking and writing good English. For the essence of good mathematics, like good language, is, first to know what you want to say, and then to be able to frame words which express exactly that and nothing else."

— E. Cunningham, *Eureka*, no. 5 (1941), p. 19.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN
with the assistance of the mathematics departments of St. Olaf and Carleton Colleges
COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, CARLETON COLLEGE

We invite readers to submit reviews of significant recent college-level mathematics books. We especially encourage reviews based on classroom use, or comparative reviews of several related books. Reviews should ordinarily not exceed two pages (per book) typed double spaced. Manuscripts of reviews as well as books submitted for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057.

Elementary Differential Equations and Boundary Value Problems. Second edition. By William E. Boyce and Richard C. DiPrima. Wiley, New York, 1965, 1969. xiv + 533 pp. (appendix and index) \$16.25. (Telegraphic Review, October 1969.)

Differential Equations and Their Applications, An Introduction to Applied Mathematics. By Martin Braun. Springer-Verlag, New York, 1975. xiv + 718 pp. \$14.80 (paper). (Telegraphic Review, December 1975.)

I have used both texts with success in a course populated by junior, senior, and (a very few) sophomore students. These students are largely majors in mathematics, chemistry, or physics; occasionally we attract a student from biology or economics. Of the two texts Boyce and DiPrima is the more traditional and the more thorough. Both are remarkably free of errors. Boyce and DiPrima include answers for all exercises while Braun gives no answers to exercises.

Both texts treat the major varieties of first order equations for which analytic means of solution are known (i.e. linear, variables separable, and exact equations). Boyce and DiPrima have a section on integrating factors while Braun hardly touches integrating factors at all. Boyce and DiPrima also use their exercises to cover additional varieties of first order equations such as the Bernoulli and Clairaut equations. Both texts wisely introduce existence and uniqueness theorems only after the student has solved a great many first order equations. Braun follows his presentation of the existence theorem with two sections on solution of the equations $f(x) = x$ and $g(x) = 0$ by iteration. These sections can be very helpful. My students, for example, were unfamiliar with iteration techniques and thus found the proof of the existence theorem very strange.

Braun ends his treatment of first order equations with a section on numerical techniques. Boyce and DiPrima, on the other hand, consider numerical techniques much further along. Typically their coverage is more complete, including predictor corrector methods and numerical techniques for systems of first order equations; neither topic appears in Braun.

In the sections on linear second order equations Braun's examples are superb. For instance, in addition to the usual mechanical vibrations and electrical networks he includes an excellent discussion of a model for the detection of diabetes. Boyce and DiPrima include a section on Euler equations and use these equations to motivate the study of regular singular points and the method of Frobenius. Braun mentions Euler equations only in a few exercises and his treatment of regular singular points seems to "come out of the air" as my students were quick to note.

The two texts have comparable treatments of the Laplace transform although I was troubled by the fact that Braun nowhere lists a table of transforms. Both texts treat equations with discontinuous "forcing functions" but Braun does the better job of motivating the Dirac delta function. He also gives a brief outline of Laurent Schwartz's treatment of the delta function. (This apparently was helpful only to those students who had previously studied abstract algebra.) Incidentally Boyce and DiPrima have 5 sections on higher order equations and Braun has none.

Both texts give a standard treatment of systems of equations but Braun handles the computational

necessities of systems of equations with more ease than Boyce and DiPrima. In particular the technique he uses to solve $x' = Ax$ when A has an eigenvalue of multiplicity 2 or more is much less messy than that given in Boyce and DiPrima. Also Boyce and DiPrima mention e^{At} only in one exercise; this of course means that their solution of $x' = Ax + f$ by variation of parameters is unnecessarily complex.

Braun does a good job on stability of linear systems and stability of equilibrium solutions; his treatment is in the function space style. Braun's examples also are characteristically excellent; those of Boyce and DiPrima are virtually non-existent. Boyce and DiPrima, however, cover the Liapounov functions for systems of two equations while Braun does not mention Liapounov functions.

Both texts treat the heat and wave equations in a section on Fourier series and separation of variables. One should also note that Boyce and DiPrima include an expanded treatment of boundary value problems and the Sturm–Liouville theory. There is also an edition of Boyce and DiPrima which ends with separable partial differential equations.

D. F. BAILEY, Cornell College

Statistical Distributions. A Handbook for Students and Practitioners. By N. A. J. Hastings and J. B. Peacock. Halsted Press, New York, 1975. xiii + 130 pp. \$5.95. (Telegraphic Review, June-July 1975.)

In a first course in Mathematical Statistics the student is confronted with a wide variety of distributions. This handbook provides the student, practitioner, or teacher with an attractive summary of twenty-five distributions.

Let us consider the authors' summary of the normal distribution. We find that they list the range of the variate, the probability density function, moment generating function, characteristic function, mean, r^{th} moment, variance, mode, median, coefficients of skewness and kurtosis, and other details. Then they tabulate parameters, their sample estimators, and properties of these estimators (e.g. unbiased or maximum likelihood). In a few sentences they summarize facts concerning the distribution of linear combinations of independent normal variates, and of sums of squares of such variates. Finally, and this I find very useful, they describe a method for generating samples from a normal population using a computer.

This last information is useful in the classroom where large amounts of data are not readily available for analysis. One unsatisfactory aspect of traditional statistics courses is that the student had little experience in handling a large mass of data. Now it is easy for the teacher to generate a large sample from a particular distribution on a computer, and ask the students to identify the parent distribution. Such a project calls on the students to use many of the theoretical aspects of the course and to face the practical problems associated with handling large samples. This handbook will assist teachers to design such projects and assist students to complete them.

Finally, a few remarks on the presentation of the text are in order. The text is necessarily terse, like all good summaries. It makes an excellent complement to a standard text such as J. Freund's *Mathematical Statistics.* The distributions are treated alphabetically: this is quite helpful in "getting about" the text. The book is well bound with a hard cover.

Certainly teachers will benefit from having a copy close at hand. Students who are going to study statistics for a year or more will find the handbook a useful inexpensive summary of facts.

T. M. MILLS, Bendigo College of Advanced Education, Victoria, Australia

TELEGRAPHIC REVIEWS

Telegraphic reviews are designed to give prompt notice of new books with sufficient information to assist our readers in deciding whether to order an examination copy or to suggest library purchase. Possible uses are indicated as follows:

T = textbook P = professional reading
S = supplementary reading L = undergraduate library purchase
13 to 18 = freshman to second year graduate level usage
1 to 4 = appropriate time in semesters to cover text

Asterisks (*) or question marks (?) denote special positive or negative emphasis, respectively. Publishers are denoted by standard abbreviations; complete addresses may be found in *Books in Print*.

GENERAL, S(13), L*. *Inequalities*. P.P. Korovkin. Trans: Sergei Vrubel. MIR (Imported by: Imported Pub, 320 W. Ohio St., Chicago, IL 60610), 1975, 72 pp, $1 (P). A delightful entry in the Little Mathematics Library. 35 problems and their solutions comprise an introduction to certain inequalities and their uses, and 28 more exercises challenge the reader, who needs only to know precalculus mathematics. Endless fascination! DFA

GENERAL, S?, P?. *Calculator Users Guide and Dictionary*. Charles J. Sippl. Matrix Pub, 1976, xv + 428 pp, $9.95 (P). Guide covers all types of calculators, but concentrates on programmable ones. Offers criteria for comparison and evaluation, while describing some calculators in detail. Descriptions vary greatly in style and seem to correspond most closely to manufacturer's blurbs. Main source for the dictionary portion is the author's *Computer Dictionary and Handbook* (TR, March 1969). Definitions tend to be imprecise and confusing, if not actually incorrect. RSK

GENERAL, S, L. *1975 U.S. Computer Chess Championship*. David Levy. Comp Sci Pr, 1976, v + 86 pp, $6.95; $4.95 (P). Brief explanation of how computers play chess; short history of computer chess; analysis and description of all tournament games (twelve machines). LCL

BASIC, T(13). *The Power of Relevant Mathematics: Basic Concepts*. Kenneth L. Whipkey, Mary Nell Whipkey, Joanne Jarocki. P-H, 1977, xvii + 521 pp, $12.95. Topics in secondary school mathematics--arithmetic, geometry, linear and quadratic equations, plus material on statistics and Basic. First chapter introduces students to algorithms and flowcharts, which are used throughout the text (e.g., to show how to solve two equations in two unknowns). "Real-life" applications in exercises; suggested readings follow each chapter. JG

PRECALCULUS, T(13: 1, 2). *Principles of Mathematics, Third Edition*. Paul K. Rees, Charles Sparks Rees. P-H, 1977, xiv + 494 pp, $13.95. Standard high-school mathematics topics. Chapter on calculus uses limit notation, but does not discuss limits per se. Includes problems for hand-held calculators. (TR, *Second Edition*, February 1972.) JG

PRECALCULUS, T(13), S*. *Graphing Functions, Module XI, Series in Mathematics Modules*. Gary Fitts. Cummings, 1977, 167 pp, $3.50 (P). Eight self-study lessons include discussion of domain, asymptotes (symmetry not included). Sets stage for further development by using polynomials (degree ≤ 3) to discuss maximum (minimum) value and increasing (decreasing) intervals, slope (between two points on a curve) and rectangular approximation to area. LCL

PRECALCULUS, T(13), S*. *Functions and Word Problems, Module X, Series in Mathematics Modules*. Gary Fitts. Cummings, 1976, iv + 123 pp, $2.35 (P). Eight self-study lessons focusing on functional notation and translating English sentences into mathematical formulas. LCL

EDUCATION, T(13-14: 1), L. *Patterns and Systems of Elementary Mathematics*. Jonathan Knaupp, *et al*. HM, 1977, xiv + 425 pp, $12.95 (P). An activity-based workbook of elementary school topics designed for prospective teachers and non-science majors. Emphasizes use of puzzles, games and projects to develop a hands-on learning environment. LAS

HISTORY, P. *Elliptic Functions According to Eisenstein and Kronecker*. André Weil. Ergebnisse der Math., B. 88. Springer-Verlag, 1976, 92 pp, $14.80. An introduction to elliptic function through the eyes of Eisenstein and Kronecker. The author describes, interprets, and amplifies the work of these mathematicians. He shows that the papers of Eisenstein and Kronecker contain a treasure of rich ideas for modern mathematicians. A valuable book for historians, number theorists, and function theorists. SG

HISTORY, P. *Bernard Bolzano*. Jan Berg. Friedrich Frommann Verlag, 1977, 234 pp. A volume of the collected works, containing the texts of previously unpublished writings on mathematics and philosophy from the period 1810-16. JD-B

FOUNDATIONS, P. *Generalized Recursion Theory, Proceedings of the 1972 Oslo Symposium*. Ed: J.E. Fenstad, P.G. Hinman. Stud. in Logic and Found. of Math., V. 79. North-Holland, 1974, viii + 456 pp, $34.75. Seventeen papers presenting a broad view of methods and results--the only comprehensive presentation of the subject in book form. Extensive and useful bibliography. LCL

FOUNDATIONS, T(13-15), S, L. *Elementary Set Theory, Proof Techniques*. Carl E. Gordon, Neil Hindman. Hafner Pr, 1975, xi + 305 pp, $8.95. Propositional calculus and first order logic followed by rather formal, meticulous development of set properties: intersections, unions, relations, functions and mathematical induction. Emphasis on proof techniques. Classes are not mentioned, except indirectly in an optional section on Russell's paradox. LCL

NUMBER THEORY, S(18), P. *Lecture Notes in Mathematics-555: The Genus Fields of Algebraic Number Fields*. Makoto Ishida. Springer-Verlag, 1976, v + 115 pp, $7.40 (P). Notes taken from lectures given by the author which include recent results about genus fields of algebraic number fields. CEC

NUMBER THEORY, S(18), P. *A Theoretical and Computational Study of Generalized Aliquot Sequences.* H.J.J. Te Riele. Math. Centre Tracts, No. 74. Math Centrum, 1976, x + 76 pp, Dfl. 10 (P). Generalized aliquot sequences are sequences every term of which is the sum of certain, but not necessarily all, aliquot parts of the preceding term. This monograph presents a theoretical and computational study of these sequences. CEC

LINEAR ALGEBRA, T*(13-14: 1), S, L. *Elementary Matrix Algebra with Applications, Second Edition.* Richard J. Painter, Richard P. Yantis. Prindle, 1977, viii + 424 pp, $14.95. A solid beginning course in linear algebra which includes applications from acturarial science, business, economics, engineering, manpower planning, operations research, psychology, statistics, and transportation science. Lots of good examples and problems. CEC

ALGEBRA, T(15-16: 1), L. *Abstract Algebra: A First Look.* Joseph E. Kuczkowski, Judith L. Gersting. Pure and Appl. Math., V. 38. Dekker, 1977, viii + 323 pp, $17.50. Organized from the point of view of category theory: main topics of discussion are structures, substructures, morphisms, quotient structures, with groups, rings, vector spaces studied as detailed examples. Easy exercises scattered throughout the exposition, as well as many exercises at the end of each chapter. Clearly written. JG

FINITE MATHEMATICS, T(13: 1, 2). *Finite Mathematics: A Modeling Approach.* Marvin L. Bittinger, J. Conrad Crown. A-W, 1977, xi + 443 pp, $13.95. Usual topics, each with applications, laid out in versatile modular form. Careful buildup with plenty of routine work. Requires neither computer nor hand calculator. LCL

FINITE MATHEMATICS, T(13: 1). *Finite Mathematics.* F. Lane Hardy. Har-Row, 1977, xiv + 448 pp, $11.95. The usual topics (probability and statistics, matrices, linear programming) for students with limited mathematics background. Plenty of problems. Appendix with computer programs (Basic). JG

CALCULUS, T*(13: 2, 3), L. *Calculus and Analytic Geometry.* Al Shenk. Goodyear, 1977, xiv + 893 pp, $19.95. Calculus of one, two, and three variables, including introduction to vector analysis and to differential equations. Emphasis on use of geometric intuition. More difficult theoretical topics covered in truly optional sections. Outstanding exercises. Applications (treated primarily in the exercises) are drawn from economics, social science, biology, chemistry, forestry, navigation, and many others. Historical notes in optional sections are welcome; unfortunately, there are only three. Dizzying variety of typefaces. JG

CALCULUS, T(13: 1, 2). *Calculus, A Tool for Analysis and Decision.* Michael L. Kovacic. Prindle, 1977, xii + 395 pp, $13.95. Calculus with emphasis on applications to business, economics and other social sciences, life sciences. Intuitive, with few proofs. Over 2000 problems. Includes techniques for finding extrema of functions of several variables. Good, but sometimes careless ("$\{x|x\varepsilon R\}$" is an "important equation"; elsewhere, "2+3=$5"). Reviews algebra and trigonometry. DFA

CALCULUS, T?(13: 2). *Elementary Calculus.* T.M. Cronin. Bradford U Pr, 1976, vii + 206 pp, £3.25 (P). Integration, differentiation, their relationship, applications of both. Much too terse, not enough exercises. (Examples: the rectangle, trapezoidal, tangent, and Simpson rules, with error estimates, are presented in 4 pages; so is the material on max-min problems, including exercises.) No treatment of logarithmic or exponential functions (but "ln x" and "e^x" appear in isolated examples). Too brief for American use. DFA

CALCULUS, T(13: 2). *Calculus with Analytic Geometry, A First Course, Third Edition.* Murray H. Protter, Charles B. Morrey, Jr. A-W, 1977, ix + 689 pp, $15.95. Revisions from second edition (TR, August 1970) include a reworking of chapters on vectors, i.e., Chapters 12 and 16 have become Chapter 13. Exercises have been added throughout and all exercises have been converted to metric units. LLK

CALCULUS, T(13: 1). *Essential Calculus with Applications.* Richard A. Silverman. Saunders, 1977, xii + 267 pp, $10.95; *Instructor's Manual,* 101 pp, (P). Concise treatment of basic topics. Introduction of ε,δ language in examples. Instructor's manual has supplementary problems and answers. LLK

CALCULUS, T*(13: 1, 2), L. *A Short Calculus, An Applied Approach, Second Edition.* Daniel Saltz. Goodyear, 1977, ix + 470 pp, $15.95. A revision of the earlier 1973 edition (TR, June-July 1973). Improvements: more and varied examples, additional exercises at all levels, motivational material to introduce new ideas. A most appealing book due to its flexibility (suited for a semester or a year) and its wide variety of examples from biology, economics, physics and psychology. The treatment is intuitive but honest. TAV

CALCULUS, T(13: 1). *Calculus for the Social Sciences.* A.W. Goodman. Saunders, 1977, xx + 442 pp, $12.95. Good presentation of topics with applications for a short course. LLK

CALCULUS, T(13: 1). *Introductory Calculus with Applications, Second Edition.* J.S. Ratti, M.N. Manougian. HM, 1977, x + 476 pp, $13.50; *Solutions Manual,* 217 pp, (P). Intuitive text for non-major calculus course. Applications to ecology, economics, biological and physical sciences, and geometry. Revised chapters on precalculus and limits. New chapter on differential equations. Exercises for pocket calculators. Solutions manual available. (First edition, TR, November 1973.) RBK

REAL ANALYSIS, T(15). *Mathematical Analysis, A Straightforward Approach.* K.G. Binmore. Cambridge U Pr, 1977, x + 257 pp, $19.95; $7.95 (P). A basic advanced-calculus-for-mathematicians text. Covers functions of one variable only. Major topics: the real number system, continuity, differentiation, Riemann integration, transcendental and gamma functions. Many exercises. SG

DIFFERENTIAL EQUATIONS, P. *Singular and Degenerate Cauchy Problems.* R.W. Carroll, R.E. Showalter. Acad Pr, 1976, viii + 333 pp, $14.50. An entry in the Mathematics in Science and Engineering series. Major topics: singular partial differential equations of Euler-Poisson-Darboux type, canonical sequences of singular Cauchy problems, degenerate equations with operator coefficients. Concern is with well-posed problems. Over 650 references. DFA

DIFFERENTIAL EQUATIONS, T(14: 1), S, L. *Schaum's Outline of Modern Introductory Differential Equations with Laplace Transforms, Numerical Methods, Matrix Methods, Eigenvalue Problems.* Richard Bronson. McGraw, 1973, 306 pp, $4.50 (P). An introduction to ordinary differential equations which includes Laplace transforms, numerical methods, matrix methods and eigenvalue problems along with the standard applications. The format is vintage Schaum Outline. CEC

NUMERICAL ANALYSIS, S(17), P. *Lecture Notes in Economics and Mathematical Systems-134: Computing Methods in Applied Sciences and Engineering.* Ed: R. Glowinski, J.L. Lions. Springer-Verlag, 1976, viii + 390 pp, $15.20 (P). 22 papers from an international colloquium held in Versailles in December 1975. About half are in English and half in French. On numerical methods and their applications to scientific problems. Linear and nonlinear systems, finite elements, dynamical problems, inverse problems, and integral methods. RWN

FUNCTIONAL ANALYSIS, T(15: 1), L. *Integral Equations: A Short Course.* Ll. G. Chambers. Intern'l Textbooks, 1976, 198 pp, $17.50. Introductory, providing motivation for studying the subject. Chapters on Fredholm equations, Volterra equations, the use of various transforms, and approximate methods. Many examples and exercises. Assumes familiarity with matrices. An attractive book which appears to be quite usable by undergraduates. DFA

OPTIMIZATION, T(16-17: 1), P. *Optimisation, Méthodes Numériques.* A. Auslender. Masson (U.S. Distr: SMPF, 111 W. 57th St., NY 10019), 1976, 178 pp, $29.90. Mathematical aspects of numerical methods for optimization. Covers linear programming, duality, descent methods, convex nondifferentiable programming, decomposition methods and the use of variational inequalities. Bibliography and references to applications. RWN

OPTIMIZATION, P. *Lecture Notes in Economics and Mathematical Systems-124: The Computation of Fixed Points and Applications.* Michael J. Todd. Springer-Verlag, 1976, vii + 129 pp, $7.40 (P). Brouwer's theorem and Kakutani's theorem and their extensions. Approximation methods based on triangulations and simplices. Restart and homotopy algorithms. Applications to economic equilibria and optimization. Presumes real analysis. RWN

ANALYSIS, P. *Locally Convex Algebras in Spectral Theory and Eigenfunction Expansions.* H.G.J. Pijls. Math. Centre Tracts, No. 66. Math Centrum, 1976, xi + 97 pp, Dfl. 12 (P).

ANALYSIS, T(15: 1), *Fourier Series.* Georgi P. Tolstov. Trans: Richard A. Silverman. Dover, 1976, x + 336 pp, $5 (P). From the second Russian edition, to which problems have been added. First published by Prentice-Hall in 1962. Later chapters discuss double Fourier series, the Fourier integral, Bessel functions, Fourier-Bessel series, applications to mathematical physics. Many derivations, some proofs, not a lot of exercises. DFA

ANALYSIS, T(18), P. *Elements of the Theory of Representations.* A.A. Kirillov. Trans: Edwin Hewitt. Grund math. Wissenschaften B. 220. Springer-Verlag, 1976, xi + 315 pp, $34.90. Basic concepts and principal methods, with a large section on the method of orbits. Illustrates the general theory with examples and includes a short historical sketch. Many problems and references. Reviews needed background from topology, module theory, functional analysis, topological group and Lie theory. DFA

ANALYSIS, T(18: 1), P. *Multiple Hypergeometric Functions and Applications.* Harold Exton. Wiley, 1976, 312 pp, $27.50. Relationship to ordinary hypergeometric functions and the Lauricella functions. Integral representations. Generating functions. Analytic continuation. Solutions to systems of ordinary differential equations. Applications to statistics and physics. RWN

DIFFERENTIAL GEOMETRY, T(16-18: 1), S, P*, L*. *Singularity Theory and an Introduction to Catastrophe Theory.* Yung-Chen Lu. Springer-Verlag, 1976, xii + 199 pp, $12 (P). A self-contained survey of the mathematics of catastrophe theory, presuming only sophisticated advanced calculus and elementary differential geometry. Contains a great number of helpful examples and intuitive remarks. Revised notes from expository lectures delivered at the Battelle Seattle Research Center in April, 1975. LAS

DIFFERENTIAL GEOMETRY, P. *Differential Geometry and Relativity.* Ed: M. Cahen, M. Flato. Math. Physics & Appl. Math., V. 3. Reidel, 1976, xi + 304 pp, $29.50. A collection of contributed articles by friends of André Lichnerowicz in honor of his sixtieth birthday. JAS

GEOMETRY, S(13-16), L. *Dividing a Segment in a Given Ratio.* N.M. Beskin. Trans: V. Zhitomirsky. MIR (Imported by: Imported Pub, 320 W. Ohio St., Chicago, IL 60610), 1975, 71 pp, $1 (P). "The most elementary problem has hidden in it unexpected connections with other problems...". Beginning with the elementary problem of dividing a line segment in a given ratio, we encounter problems in affine and projective geometry, and group theory, though no knowledge of these subjects is presupposed. An excellent example of the process of mathematical discovery. JG

GEOMETRY, T*(13-14: 1), *Experiencing Geometry.* James V. Bruni. Wadsworth, 1977, x + 310 pp, $11.95 (P). Filled with concrete models and illustrations. Forces experimentation with various geometric figures. Encourages a discover-for-yourself approach: hence, the title. No formal theorems and proofs. Very well done. Exercises (experiences). Bibliography. Index. RJA

GEOMETRY, S(13-14). *Four-Dimensional Geometry, An Introduction*. Adrien L. Hess. NCTM, 1977, iii + 28 pp, $1.60 (P). Booklet presenting the history and definition of four-dimensional geometry. Includes drawings and models and instructions for studying them. Bibliography. RJA

TOPOLOGY, S(16-17), P. *Vorlesungen über die Theorie der Polyeder*. E. Steinitz, H. Rademacher. Grund. math. Wissenschaften, B. 41. Springer-Verlag, 1976, viii + 351 pp, $34.40. A reprint of the original 1934 edition. JD-B

TOPOLOGY, T(18: 1), S. *Hausdorff Compactifications*. Richard E. Chandler. Lect. Notes in Pure and Appl. Math., V. 23. Dekker, 1976, vii + 146 pp, $16.50 (P). Imbed a space in a compact Hausdorff space and take its closure to get a compactification. This is a study of such compactifications. JAS

TOPOLOGY, P. *The Oriented Bordism of Topological Manifolds and Integrality Relations (Revised Edition)*. Ib Madsen, R. James Milgram. Aarhus U, 1976, v + 248 pp, (P). An up-to-date report on classification results for piecewise linear and topological manifolds. JAS

PROBABILITY, P. *Lecture Notes in Mathematics-550: Proceedings of the Third Japan-USSR Symposium on Probability Theory*. Ed: G. Maruyama, J.V. Prokhorov. Springer-Verlag, 1976, vi + 722 pp, $21 (P). Proceedings from the symposium held at Tashkent, USSR from August 27 to September 2, 1975. JAS

PROBABILITY, P. *Lecture Notes in Economics and Mathematical Systems-121: On Regenerative Processes in Queueing Theory*. J.W. Cohen. Springer-Verlag, 1976, ix + 93 pp, $7.40 (P). Exploits the rather recent formulation of a "stochastic mean value theorem" for regenerative processes to provide new insights into well known relationships in queueing theory. An interesting and fruitful approach. TAV

PROBABILITY, T?(18: 2), P. *Probability Theory on Boolean Algebras of Events*. O. Onicescu, I. Cuculescu. Editura Academiei (Romania), 1976, 185 pp, Lei 25. In the preface the authors note that there is no unified treatment of this subject where a beginner can learn the theory, and offer this treatise to meet that need. It will be tough sledding for the "beginner" still, for this treatment requires substantial sophistication, e.g., facility with lattice theory, tensor products, and free algebras. The bibliography contains numerous references, most written in Romanian. TAV

PROBABILITY, S*(15-16), L. *The Monte Carlo Method*. I.M. Sobol. Trans: V.I. Kisin. MIR (Imported by: Imported Pub, 320 W. Ohio St., Chicago, IL 60610), 1975, 73 pp, $1 (P). A different translation of a brief monograph first published in the U.S. in 1974 by U. of Chicago Pr. (TR, February 1976). This new version, published in Russia, is one-third the cost of the U.S. version. LAS

PROBABILITY, P. *Contracting Markov Decision Processes*. J.A.E.E. van Nunen. Math. Centre Tracts, No. 71. Math Centrum, 1976, vii + 137 pp, Dfl. 17 (P). Markov reward processes, decision processes; stopping times and contraction; value oriented successive approximation; upper, lower bounds and suboptimality; applications to solving linear systems and an inventory problem; references; index. RBK

PROBABILITY, P. *Lecture Notes in Mathematics-547: Measures on Topological Semigroups: Convolution Products and Random Walks*. Arunava Mukherjea, Nicolas A. Tserpes. Springer-Verlag, 1976, v + 197 pp, $10.20 (P). A development of the theory of random walks on topological semi-groups. JAS

PROBABILITY, P. *Lecture Notes in Mathematics-566: Empirical Distributions and Processes*. Ed: P. Gaenssler, P. Révész. Springer-Verlag, 1976, 146 pp, $8 (P). Selected papers from the meeting at Oberwolfach, March 28 to April 3, 1976. Papers were selected to conform to the title. JAS

PROBABILITY, T(18), P. *Statistical Theory of Reliability and Life Testing, Probability Models*. Richard E. Barlow, Frank Proschan. HR&W, 1975, xiii + 290 pp, $19.95. In the International Series in Decision Processes and the Series in Quantitative Methods for Decision Making. First of two books--this one is probabilistic and emphasizes the newer, research aspects of reliability theory, while the second will deal with inferential aspects of reliability and life testing. Treats a number of new classes of life distributions and considers various types of positive dependence among random variables. RSK

PROBABILITY, T(15-16: 1, 2). *Probability Theory with the Essential Analysis*. J. Susan Milton, Chris P. Tsokos. Appl. Math. and Comp., No. 10. A-W, 1976, xix + 339 pp, $14.50 (P); $24.50. Axiomatic approach to probability using, but not presupposing, elementary measure theory. Topics covered include n-dimensional random variables, modes of convergence, limit theorems. References are included at the end of each chapter. Few problems. JG

STATISTICS, S*(13-16), L*. *Statistics: A Guide to the Study of the Biological and Health Sciences*. Ed: Judith M. Tanur, *et al.* Holden-Day, 1977, ix + 140 pp, $3.50 (P). Eleven essays taken from the editors' well-known 1972 collection *Statistics: A Guide to the Unknown* (TR, January 1973; extended review, April 1974), plus one new essay, describing applications of statistics to the biological and health sciences. Second of the "mini-SAGTU's" to appear, the first being *Statistics: A Guide to Business and Economics* (TR, November 1976). RSK

STATISTICS, P. *Statistical Methods in Research and Production with Special Reference to the Chemical Industry, Fourth Revised Edition*. Ed: Owen L. Davies, Peter L. Goldsmith. Longman, 1976, xiii + 478 pp, $14.50. A compendium of statistical techniques from histograms to computer-assisted multiple regression with applications chosen exclusively from the chemical industry. A useful reference work for the researcher or engineer. TAV

STATISTICS, T?(13-14: 1, 2). *Statistics for Business and Economics, Third Edition*. Stephen P. Shao. Merrill, 1976, xiii + 815 pp, $15.50. Up-dated version of the 1972 *Second Edition* (TR, April 1973). A confusing text, with an incredible amount of material on descriptive statistics, much of which is outdated, and a strong emphasis on time series analysis. RSK

STATISTICS, T(13-14: 1, 2). *Statistical Methods for Engineers and Scientists.* Robert M. Bethea, Benjamin S. Duran, Thomas L. Boullion. Statistics, V. 15. Dekker, 1975, xxi + 583 pp, $25.50. Standard topics plus chapters on orthogonal polynomials for regression, and experimental design. No proofs. Worked examples for each method. Problems and examples mostly from chemical engineering. Small data sets provided for many problems. RBK

STATISTICS, P. *Compstat 1976: Proceedings in Computational Statistics.* Ed: Johannes Gordesch, Peter Naeve. Physica-Verlag, 1976, 496 pp, DM 56 (P). Proceedings of the second Compstat Symposium, held in West Berlin in 1976. Contains 58 short papers in the following subject areas: computational probability, cluster analysis and multidimensional scaling, numerical and algorithmic aspects of statistical models, simulation and stochastic processes, software, and applications. RSK

STATISTICS, S. *Mathematical, Statistical and Financial Tables for the Social Sciences.* Z.W. Kmietowicz, Y. Yannoulis. Longman, 1976, ix + 54 pp, $2.95 (P). Brief collection of tables, primarily statistical. Mathematical section includes factorials and binomial coefficients, in addition to some standard tables. Financial section consists of compound interest and annuity tables. Statistical tables are in four subsections: basic probability distributions, correlation tests, nonparametric tests, and random numbers. RSK

STATISTICS, P**, L*. *Handbook of Statistical Distributions.* Jagdish K. Patel, C.H. Kapadia, D.B. Owen. Statistics, V. 20. Dekker, 1976, xiv + 302 pp, $19.75. A collection of useful results, concisely presented, with references to readily available sources. Chapter topics are: moments, cumulants, and generating functions; inequalities; order statistics; families of distributions; characterization of distributions; point estimation; confidence intervals; properties of distributions; basic limit theorems; miscellaneous topics. RSK

STATISTICS, T?(14-16: 2). *Introduction to Statistics and Computer Programming, Pilot Edition.* Carl F. Kossack, Claudia I. Henschke. Holden-Day, 1975, 651 pp, $10.95 (P). Attempt to integrate statistics and computer programming for students in the social, biological, management and behavioral sciences, with an emphasis on decision theory. Presumes only high school algebra, but introduces complicated functions, matrices and integration from a numerical evaluation point of view. Although there are some good computer programs (in Fortran) to illustrate statistical ideas, the coverage of statistical topics is very uneven and at times the presentation is overwhelming. RSK

STATISTICS, T?(13: 1). *Elementary Computer-Assisted Statistics.* Frank Scalzo, Rowland Hughes. Petrocelli/Charter, 1975, xv + 345 pp, $13.95. A beginning chapter on programming in Basic, with thirteen simple programs scattered throughout the text (mean, variance, 90% confidence intervals, regression line, etc.). Minimal statistics coverage with discussion limited to techniques only. LCL

STATISTICS, T(14-15: 1, 2). *Introductory Statistical Analysis.* Donald L. Harnett, James L. Murphy. A-W, 1975, x + 565 pp, $14.95. Designed for business, economics or social science students, it presumes no calculus yet requires a fair amount of mathematical sophistication. In addition to the standard topics, it includes chapters on statistical decision theory, multiple regression, and time-series and index numbers, and incorporates nonparametric methods and analysis of variance into other chapters. The presentations are terse and the reading heavy for a course at this level. RSK

STATISTICS, T(18), P. *Time Series, Data Analysis and Theory.* David R. Brillinger. HR&W, 1975, xii + 500 pp, $19.95. In the International Series in Decision Processes. First of two volumes--this one is devoted to aspects of the linear analysis of stationary vector-valued time series, while the second will deal with nonlinear analysis and extensions to stationary vector-valued continuous series, spatial series, and vector-valued point processes. Basic tool employed in the analysis of time series is the finite Fourier transform. Proofs are separated from the textual material and placed at the end of the book. Good set of references. RSK

STATISTICS, T?(14-17: 1, 2), S. *Probability; Decision; Statistics.* James V. Bradley. P-H, 1976, xviii + 586 pp, $18.50. Assumes only high school algebra, but would require much mathematical sophistication, particularly in the later chapters. Contains some excellent sections, but the overall balance is strange. Probability and decision theory are given detailed coverage, but statistics is not (e.g., there is nothing on regression and correlation). Has a good emphasis on assumptions and appropriateness of procedures. On the other hand, 50 pages are used to describe and illustrate the "General Limit Effect", while in comparison about 40 pages are used to cover all confidence intervals and hypothesis tests involving the normal and t distributions. RSK

STATISTICS, T(16-17: 2), P. *Statistical Models in Applied Science.* Karl V. Bury. Wiley, 1975, xvii + 625 pp, $24.95. In the Wiley Series in Probability and Mathematical Statistics. Designed for engineers and applied science students, it concentrates on relevant mathematical statistics topics. Divided into three parts: Part I presents results and methods of general statistical analysis; Part II discusses appropriate continuous statistical models: normal, log-normal, gamma, beta, type I extreme value, and Weibull; Part III provides four applications of these models: probabilistic design, component reliability, system analysis, and material strength. No exercises. RSK

Reviewers Whose Initials Appear Above

Richard J. Allen, St. Olaf; David F. Appleyard, Carleton; Clifton E. Corzatt, St. Olaf; John Dyer-Bennet, Carleton; Jennifer Galovich, St. Olaf; Steven Galovich, Carleton; Lorraine L. Keller, St. Olaf; Roger B. Kirchner, Carleton; Richard S. Kleber, St. Olaf; Loren C. Larson, St. Olaf; R.W. Nau, Carleton; J. Arthur Seebach, Jr., St. Olaf; Lynn Arthur Steen, St. Olaf; T.A. Vessey, St. Olaf.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY AT BUFFALO

Readers are invited to contribute to the general interest of this department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, NW, Washington, D.C. 20036. Items must be submitted at least five months before publication can take place.

PERSONAL ITEMS

Furman University: Professor C. R. Wylie has retired as Chairman of the Department of Mathematics; Professor D. H. Clanton has been appointed Chairman of the Department of Mathematics; Drs. P. M. Cook II and D. F. Rall have been appointed Assistant Professors.

Dr. W. E. Mainville, Jr., Associate Professor of Mathematics and Assistant Dean, College of Arts and Sciences, University of Maine at Portland-Gorham, has been promoted to Professor of Mathematics.

Dr. Haim Reingold, Professor of Mathematics and Chairman of the Department of Mathematics at the Illinois Institute of Technology, has retired with the title of Professor Emeritus; Dr. Reingold has been appointed Professor at Indiana University Northwest.

Mr. David G. Halmstad, Guardian Life Insurance of America, Ridgefield, Connecticut, died on February 19, 1977, at the age of 39. He was a member of the Association for eighteen years.

Dr. Floyd S. Harper, Lincoln, Nebraska, died on March 19, 1977, at the age of 81. He was a member of the Association for thirty-six years.

Dean Emeritus and Professor Emeritus Gillie A. Larew, Randolph Macon Woman's College, Lynchburg, Virginia, died on January 2, 1977, at the age of 94. Professor Larew was a Charter Member of the Association.

Mr. Edson C. Lockwood, Milton, Massachusetts, died on May 4, 1976. He was a member of the Association for fifty years.

Associate Professor Paul R. Schrodt, University of Louisville, died on November 9, 1976, at the age of 50. He was a member of the Association for eleven years.

Professor Luke Zaccaro, Youngstown State University, died on March 19, 1977, at the age of 53. He was a member of the Association for twenty-two years.

AMS-MAA-SIAM CONGRESSIONAL SCIENCE FELLOWSHIP FOR 1978–79

Applications are invited from candidates in the mathematical sciences for a Congressional Science Fellowship to be supported jointly by the American Mathematical Society, the Mathematical Association of America and the Society for Industrial and Applied Mathematics for the twelve-month period beginning 1 September 1978. The AMS-MAA-SIAM Fellow will serve, along with three or four Fellows selected by the American Association for the Advancement of Science and around a dozen Fellows sponsored by other scientific societies, under an annual program coordinated by the AAAS. The stipend for the 1978–79 AMS-MAA-SIAM Fellowship is $17,000, which may be supplemented by a small amount toward relocation and travel expenses. It may also be supplemented by sabbatical salary or other employer contribution in the case of a person on sabbatical leave for the 1978–79 year.

The 7 January 1977 issue of *Science*, page 55, gives a brief description of the overall AAAS program for 1976–77. As indicated there, Congressional Science Fellows spend their fellowship year working on the staff of an individual congressman or a congressional committee or in the congressional Office of Technology Assessment, the objective of the program being to enhance science-government interaction, the effective use of science in government, and the training of persons with scientific background for careers involving such use. Based on information on available congressional staff positions gathered by the AAAS during the summer, each Fellow's assignment is worked out by the Fellow and the congressional office concerned following an intensive two-week orientation and interview procedure organized by the AAAS during which the Fellows encounter many facets of Congress, the Executive Branch, and people and organizations on the Washington scene. The AAAS provides advice and assistance during this process and remains in frequent and regular contact with all the Fellows throughout the fellowship year. More detailed information about the program as a whole may be requested from Dr. Richard Scribner, Director, AAAS Congressional Science Fellow Program, 1776 Massachusetts Ave., N.W., Washington, D.C. 20036; telephone (202) 467-4475.

The AMS-MAA-SIAM Congressional Science Fellowship is to be awarded competitively to a mathematically trained person at the postdoctoral to mid-career level without regard to sex, race, or ethnic group. Selection will be made by a panel of the AMS-MAA-SIAM Joint Projects Committee for Mathematics, a nine-member committee consisting of three representatives from each of these organizations, with the cooperation and advice of Dr. Scribner. *Applications should be sent to the Conference Board of the Mathematical Sciences, 2100 Pennsylvania Ave., N.W., #832, Washington, D.C. 20037. The deadline for receipt of applications is 15 February 1978,* and it is anticipated that the award will be made by around 1 April 1978.

In addition to demonstrating exceptional competence in some areas of the mathematical sciences, an applicant for the AMS-MAA-SIAM Congressional Science Fellowship should have a rather broad scientific and technical background and a strong interest in the uses of the mathematical and other sciences in the solution of societal problems. He or she should also be articulate, literate, flexible and able to work effectively with a wide variety of people. An application should state why the applicant wants to be a Congressional Science Fellow, should summarize his or her qualifications, and should be accompanied by a resume. Also, CBMS should receive by 15 February 1978 three letters from knowledgeable persons about the applicant's competence and suitability for the award.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

APRIL MEETING OF THE NORTH CENTRAL SECTION

The 1977 spring meeting of the North Central Section of the MAA was held at North Hennepin Community College, Minneapolis, on April 29–30, 1977. There were 137 persons in attendance, including 89 MAA members and 33 students.

At the business meeting some minor revisions of the section by-laws were passed, recognition was given to top finishers and their advisors for this year's Putnam competition, reports concerning high school mathematics contests in Minnesota and North Dakota were presented, and the following elections were completed: Chairman-elect, Joseph Konhauser, Macalester College, St. Paul, three-year term on the Executive Committee; Secretary-Treasurer, Steven Galovich, Carleton College, Northfield, one-year term; Members-at-large of the Executive Committee, James Baglio, North Hennepin Community College, and Virginia Christian, Mankato State University, Mankato, one-year terms.

The principal speaker was Calvin Lathan, Monroe Community College, Rochester, Editor of The Two-Year College Mathematics Journal. He brought greetings from the national MAA office, spoke briefly about the development of the TYCMJ, and then spoke on the topic of *Small Group Instruction.*

Jay Goldman, University of Minnesota, Minneapolis, was the invited speaker for the Friday evening session. The title of Professor Goldman's presentation was *The Combinatorial Way of Thinking.*

The Saturday morning session included a panel discussion on *The Mathematics Curriculum of the First Two Years of College.* Panelists were Ken Becker, Inver Hills Community College, Inver Grove Heights; Don Marxen, Saint Mary's College, Winona; Murph Monahan, South Dakota State University, Brookings; and Steve Mondy, Anoka-Ramsey Community College, Coon Rapids.

The Saturday afternoon session consisted of the following contributed papers:

Matrices and Pythagorean triples, by Sylvan Burgstahler, University of Minnesota, Duluth.

The nonlinear recurrence $a_n = [\sqrt{3a_{n-1}(a_{n-1}+2)}]$, by Dan Amundson, South Dakota State University, Brookings.

An interesting property of ring idempotents, by Michael Larsen, St. Cloud State University, St. Cloud.

Compactifications revisited, by James Hatzenbuhler and Don Mattson, Moorhead State University, Moorhead.

Midpoint solutions of $x^x = a^b$; *A calculator inspired theorem,* by Gerald Heuer, Concordia College, Moorhead.

Gauss the Geometer, by Sy Schuster, Carleton College, Northfield.

LOUIS GUILLOU, *Secretary-Treasurer*

APRIL MEETING OF THE OHIO SECTION

The Ohio Section of the MAA held its annual Spring meeting at Denison University, Granville, Ohio, April 15 and 16, 1977. Approximately one hundred and seventy people were in attendance. Section Chairman J. A. Murtha presided; M. D. Wetzel was the Program Chairperson.

Invited addresses included: *A Survey of Recent Developments in the Representation Theory of Finite Groups*, by C. W. Curtis, University of Oregon, and currently Visiting Professor at University of Michigan; and *Designs, and Some of Their Connections to Other Branches of Mathematics*, by D. R. Hughes, University of London, and currently Visiting Professor at Ohio State University. A highlight of the meeting included a *Panel Discussion on Employment*, moderated by A. Sterrett, Denison University, with discussion leaders: L. Albers, NASA, Lewis Research Center; L. K. Barrett, University of Tennessee; J. H. Carney, Lorain County Community College; and R. E. Thomas, Battelle Institute, Columbus.

The following contributed papers were also presented:

Creativity exercises enhance research ability, by J. Altinger, Youngstown State University.

Groups of order automorphisms of partially ordered sets, by M. A. Bardwell, student, Bowling Green State University.

On tangential 2-blocks, by B. T. Datta, Ohio State University, Lima.

Determining the continuous characters of a matrix group — an application to Lie theory, by T. A. Farmer, Miami University.

Rings with fixed point fixed group actions, by J. W. Fisher, University of Cincinnati.

An introduction to coding theory, by W. Heidler, student, Miami University.

Two-element, two-operation algebras, by C. H. Heinke, Capital University.

The theology of mathematics, by R. G. Laatsch, Miami University.

Finite group actions of simple rings, by J. Osterburg, University of Cincinnati.

Random walk and Brownian motion, by C. Park, Miami University.

Orderable permutation groups, by H. L. Putt, student, Bowling Green State University.

Talent search in the inner city — a legacy of Arnold Ross, by T. G. Ralley, Ohio State University.

The cross product function in n-space, by G. L. Szoke, University of Akron.

Economic nationalization encounters optimal control theory, by W. Vaughn, student, Miami University.

H. D. Lipsich, University of Cincinnati, presented delightful 'after dinner' remarks, *It Was a Very Good Year*, honoring the 1977 bicentennial celebration of the birth of the celebrated mathematician, Carl Friedrich Gauss. Also, the agenda included the annual Business Meeting of the Section, meeting of the Executive Committee, and meeting of *ad hoc* committees: Committee on Cooperation among Colleges and Universities, Committee on Curriculum, and Committee on Teacher Training and Certification.

Section officers for academic year 1977–78 are: W. H. Beyer, University of Akron, Chairman; M. D. Wetzel, Denison University, Chairperson-Elect; J. A. Murtha, Marietta College, Past Chairman; G. Mavrigian, Youngstown State University, Secretary-Treasurer; J. H. Carney, Lorain County Community College, Program Chairman; C. A. Long, Bowling Green State University, and P. Tuchinsky, Ohio Wesleyan University, Program Committee. R. L. Wilson, Ohio Wesleyan University, serves as Sectional Governor.

GUS MAVRIGIAN, *Secretary-Treasurer*

APRIL MEETING OF THE OKLAHOMA-ARKANSAS SECTION

The thirty-ninth annual meeting of the Oklahoma-Arkansas Section of the MAA was held at Oral Roberts University, Tulsa, Oklahoma, on April 1–2, 1977. There were 79 members and 57 others in attendance, making a total of 136. Professor Verbal Snook of Oral Roberts University presided over the meeting.

At the dinner held on Friday evening in the Timko-Barton Hall at the university, Mr. M. G. Tiefenback of Hendrix College received for the second time an award for scoring the highest on the William Lowell Putman Competition among the participants from Oklahoma and Arkansas. The invited address entitled "Prime Generating Functions and Congruences" was given on Friday afternoon by Professor H. L. Alder of the University of California at Davis. Friday evening Professor Arthur Bernhart, of the University of Oklahoma, gave the fifth N. A. Court Lecture. The title was "Geometry That Counts."

The following officers were elected: Chairman, Verbal Snook, Oral Roberts University; Past Chairman, Cecil McDermott, Hendrix College; First Vice-Chairman, William Durand, Henderson State University; Second Vice-Chairman, Andrew Coe, Westark Community College, and Secretary-Treasurer E. K. McLachlan, Oklahoma State University. The following were appointed to continue as co-chairmen of the High School Mathematics contest: For Arkansas, Professor Claude Duplissey of the University of Arkansas at Little Rock; and for Oklahoma, Professor Thomas Cairns of the University of Tulsa.

In Oklahoma, 56 schools involving 2997 students and in Arkansas, 46 schools involving 2754 students participated in the high school contest.

The following papers were presented:

1. *A course in applied mathematics based on problems presented by persons from regional industries*, by Marvin Keener, Oklahoma State University.

2. *The physical problem: The design of a frequence tripler by computer simulation*, by Darrell Gimlin, Amoco Research Company.
3. *A solution based on trigonometric regression*, by Kathy Stewart, Oklahoma State University.
4. *A solution based on numerical integration*, by Jeri Ezell, Oklahoma State University.
5. *A solution based on systems of linear equations*, by Charlotte Couch and Robert Hayes, Oklahoma State University.
6. *Products of finite simple groups*, by Gary Walls, Oklahoma State University.
7. *On a class of lattice-ordered topological algebras*, by Gary Mulkey, University of Arkansas.
8. *Properties of finite quasi-injective groups*, by Dennis Bertholf, Oklahoma State University.
9. *Symmetric relations and nilpotent semigroups of set type*, by Naoki Kimura and R. B. Hora, University of Arkansas.
10. *Remarks on locally H-closed spaces*, by Larry Herrington, University of Arkansas, Pine Bluff.
11. *A product integral representation of* e^x, by William Orton, III, Hendrix College.
12. *An application of the convolution theorem to the integral of a complex function*, by Lisa Orton, Hendrix College.
13. *A Pi Mu Epsilon program*, by Emily Wonderly, Oklahoma State University.
14. *Projections onto subsets of* E^n, by Donald Hayman, Hendrix College.
15. *Modeling a glucose tolerance test*, by Jim Grooms, Oral Roberts University.
16. *Norm attaining operators on Banach spaces*, by John Wolfe, Oklahoma State University.
17. *A new class of mappings between normed spaces*, by Kiyoshi Iseki, University of Arkansas.
18. *A class of functions analytic in the unit disk*, by James Choike, Oklahoma State University.
19. *Some remarks on the lattice of convergence structures*, by Caroll Riecke, Cameron University.
20. *Norm-induced topologies*, by Janet Dillahunty, Hendrix College.
21. *A comparison of Greek and Hebrew mathematical concepts*, by Stephen Tuel, Oral Roberts University.
22. *Closures of relations*, by Alma Posey, Hendrix College.
23. *Matrix representation of finite groups*, by Michael Tiefenback, Hendrix College.
24. *R-invertible spaces*, by Paul Long, University of Arkansas.
25. *Intersection numbers*, by John Watson, Oklahoma State University.
26. *Maximizing the correlations coefficient*, by L. Franklin Kemp, Amoco Production Company.
27. *Some properties of simply presented valuated p-groups*, by Eddie Boyd, Oklahoma State University.
28. *Participation data and evaluation of the* 1972–76 *NSF Mathematics Education Program in Arkansas*, by William Orton, University of Arkansas.
29. *Metric education — The emerging national plan*, by Cecil McDermott, Hendrix College.
30. *Mathematics education in Japan*, by Kiyoshi Iseki, University of Arkansas.
31. *Teaching the foundations of mathematics using the mini computer*, by Kelvin Casebeer, Southwestern Oklahoma State University.

E. K. McLachlan, *Secretary-Treasurer*

APRIL MEETING OF THE SOUTHWESTERN SECTION

The annual meeting of the Southwestern Section of the MAA was held at Phoenix College, Phoenix, Arizona, on April 22–23, 1977. Thirty-two members and five guests registered their attendance. The following papers were contributed:

1. *Retrocirculant matrices*, by R. L. Smith, Northern Arizona University.
2. *Splitting of inseparable algebraic extensions*, by W. Y. Velez, Sandia Laboratories.
3. *Lower Riemann sums and inventory models*, by J. L. Winters, Arizona State University.
4. *Are lqrge samples ever dangerous?*, by G. S. Rogers, New Mexico State University.
5. *Counting permutations without fixed points*, by R. Bagby, New Mexico State University.
6. *On a conjecture of R. Graham*, by W. Y. Velez, Sandia Laboratories.
7. *The distribution of relatively prime integers in residue classes*, by J. E. Nyman, University of Texas at El Paso.
8. *Symmetric Pfaffians*, by A. Swimmer, Arizona State University.
9. *The homomorphism concept in high school algebra*, by E. L. Walter, Northern Arizona University.

A panel consisting of: L. Cameron, University of New Mexico; J. Geiver, New Mexico State University; R. Thompson, University of Arizona; and R. Adams, Yavapai Community College, discussed "Placement and remediation" at the Saturday morning session.

The invited speaker was J. M. Thomas, Duke University (retired). At the banquet his talk was entitled "Whither mathematics?". On Saturday morning he presented his proof of the Four-Color Theorem in a talk entitled "A brief proof." (Copies of this proof are available from Professor Thomas or the undersigned.)

Chairman of the section A. L. Seeglitz, Glendale Community College, acted as chairman of the session on contributed papers.

A. Swimmer. *Secretary-Treasurer*

APRIL MEETING OF THE TEXAS SECTION

The annual spring meeting of the Texas Section was held at Baylor University in Waco, Texas, on April 1, 2, 1977. There were 211 registered persons in attendance.

Presenting invited addresses were: Professor David Roselle, Secretary of the MAA, who spoke on "Lattice Path Problems and Combinatorial Sequences"; and Professor W. F. Lucas of Cornell University who spoke on "Modeling Coalitional Values." There was a panel discussion on "Undergraduate Mathematics Applications" presented by Professors R. M. Thrall of Rice University, E. L. Perry of Baylor University, and W. F. Lucas of Cornell University. Distinguished Service Citations for unusual contributions to mathematics and to the Texas Section were presented to Professor H. J. Ettlinger of the University of Texas and posthumously to Professor C. R. Sherer of Texas Christian University.

Officers for 1977–78 are: Chairman, H. L. Rolf, Baylor University; First Vice Chairman, R. G. Dean, Stephen F. Austin State University; Second Vice Chairman, Dalton Tarwater, Texas Tech University; Past Chairman, G. R. Blakley, Texas A & M University; Level I Director, B. D. Langston, Tarrant County Jr. College, Northeast Campus; Level II Director, Margaret Hutchinson, University of St. Thomas; Director at Large, R. H. Cranford, Texas Eastern University; Secretary-Treasurer, J. C. Bradford, Abilene Christian University; MAA High School Contest, J. R. Boone, Texas A & M University.

Contributed papers were:

Quasisimilarity and the essential spectrum, by L. R. Williams, University of Texas at Austin.

Beurling subalgebras of L^1, by G. L. Wiggins, Texas Tech University.

The trace norm on tensor products of Banach Space, by Russell Bilyeu, North Texas State University.

Methods of solution for the linear l_1 *approximation problems*, by W. Fraser, University of Texas at Austin.

Convolution operators in complex function theory, by C. N. Kellogg and R. W. Barnard, Texas Tech University.

On summation of certain infinite series, by Russell Cowan, Lamar University.

Rate invariance and Cesàro summability, by D. F. Dawson, North Texas State University.

Summability of subsequences and rearrangements of sequences, by T. A. Keagy, Wayland Baptist College.

The solution of a functional equation using transport of structure, by D. K. Cohoon, School of Aerospace Medicine and University of Texas at San Antonio.

More generalized Lipschitz conditions, by F. N. Huggins, University of Texas at Arlington.

Applications of fixed-point theory to number theory, by A. A. Mullin, Fort Hood, Texas.

Integers that are the product of $n-k$ *primes less than or equal to* 2^n, by Roy Brooks, University of Texas of the Permian Basin.

Left identities in Noetherian rings, by Baxter Johns, Baylor University.

Rings with an almost Noetherian ring of fractions, by E. P. Armendariz, University of Texas at Austin.

On the algebraic structure of pre-rings, by A. A. Mullin, Fort Hood, Texas.

On an algebraic property of ultraproducts of fields, by Joseph Heisler, St. Edward's University.

Tiling space and splitting groups, by Joe Falkner, Texas Tech University.

A note on the golden ellipse, by M. G. Monzingo, Southern Methodist University.

R_0 *spaces*, R_1 *spaces and hyperspaces*, by C. I. Dorsett, North Texas State University.

Geometric solutions of quadratic equations, by Jean Richmond, Southern Methodist University.

Alternative approach to certain geometry problems, by J. M. Stark, Lamar University.

Should linear algebra be rated R?, by Betty Barr, University of Houston.

Introducing the computer to non-science majors, by George Berzsenyi, Lamar University.

A program for recycling students in small units of foundation mathematics, by Douglas Tharp, Michael Murphy, Nancy Holder; University of Houston—Downtown Campus.

Technical mathematics course development, by Chris Boldt, Eastfield College.

Polya and the mathematically ungifted in non-technical freshman mathematics, by John Hunsucker, Sam Houston State University.

Internal variables, thermodynamic states and duality, by Vadim Komkov, Texas Tech University.

On the linearity of one-to-one entire functions, by R. K. Williams, Southern Methodist University.

Numerical quadrature: The trapezoidal rule and modifications, by Bill Copeland and Bill Anderson; East Texas State University.

Relative derivatives and chain rules, by C. R. Williams, Midwestern State University.

Some properties of valuation rings, by Nick Vaughn, North Texas State University.

On matrix inverses and sensitivity analysis, by Eugene Tidmore, Baylor University.

J. C. BRADFORD, *Secretary-Treasurer*

CALENDAR OF FUTURE MEETINGS

Sixty-first Annual Meeting, Atlanta, Georgia, January 6–8, 1978.

Fifty-eighth Summer Meeting, Brown University, August 8–10, 1978.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN, University of Pittsburgh, Pennsylvania, April 21–22, 1978, or April 28–29, 1978.

FLORIDA, St. Petersburg Junior College, Clearwater, March 3–4, 1978.

ILLINOIS, Western Illinois University, Macomb, May 5–6, 1978.

INDIANA, Indiana Central College, Indianapolis, November 5, 1977.

INTERMOUNTAIN

IOWA, University of Northern Iowa, Iowa Falls, April 22, 1978.

KANSAS, Wichita State University, Wichita, late March-early April 1978.

KENTUCKY, early April. Deadline for papers 6 wks. bef. mtg.

LOUISIANA-MISSISSIPPI, Buena Vista Hotel-Motel, Biloxi, Mississippi, February 17–18, 1978.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA, American University, Washington, D.C., November 19, 1977.

METROPOLITAN NEW YORK, late April or early May 1978. Deadline for papers 2 wks. bef. mtg.

MICHIGAN, Michigan State University, East Lansing, Spring 1978.

MISSOURI, Central Missouri State University, Warrensburg, April 7–8, 1978.

NEBRASKA, University of Nebraska at Omaha, April 14–15, 1978.

NEW JERSEY, Caldwell College, Caldwell, November 5, 1977.

NORTH CENTRAL, University of Minnesota, Morris, October 14–15, 1977.

NORTHEASTERN, Saturday after Thanksgiving, and third week in June in odd-numbered years.

NORTHERN CALIFORNIA, College of Notre Dame, Belmont, February 18, 1978.

OHIO, Wright State University, Dayton, October 28–29, 1977.

OKLAHOMA-ARKANSAS, Henderson State University, Arkadelphia, Arkansas, March 31–April 1, 1978.

PACIFIC NORTHWEST, University of Oregon, Eugene, June 16–17, 1978.

PHILADELPHIA, Moravian College, Bethlehem, November 19, 1977.

ROCKY MOUNTAIN, South Dakota School of Mines and Technology, Rapid City, April 28–29, 1978.

SEAWAY, SUNY College at Plattsburgh, New York, October 28–29, 1977.

SOUTHEASTERN, Clemson University, Clemson, South Carolina, March 31–April 1, 1978.

SOUTHERN CALIFORNIA, California State Polytechnic University, San Luis Obispo, November 11–12, 1977.

SOUTHWESTERN, New Mexico Institute of Mining and Technology, Socorro, Spring 1978.

TEXAS, Stephen F. Austin State University, Nacogdoches, March 31–April 1, 1978.

WISCONSIN, University of Wisconsin–Whitewater, late April, 1978.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Washington, February 12–17, 1978.

AMERICAN MATHEMATICAL ASSOCIATION OF TWO YEAR COLLEGES, Peachtree Plaza Hotel, Atlanta, Georgia, October 13–14, 1977.

AMERICAN MATHEMATICAL SOCIETY, Atlanta, Georgia, January 4–7, 1978.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION

ASSOCIATION FOR COMPUTING MACHINERY, Olympic Hotel, Seattle, Washington, October 17–19, 1977.

ASSOCIATION FOR SYMBOLIC LOGIC

ASSOCIATION FOR WOMEN IN MATHEMATICS, Atlanta, Georgia, January 4–8, 1978.

CANADIAN SOCIETY FOR HISTORY AND PHILOSOPHY OF MATHEMATICS/SOCIÉTÉ CANADIENNE D'HISTOIRE ET DE PHILOSOPHIE DES MATHÉMATIQUES

FIBONACCI ASSOCIATION

INSTITUTE OF MATHEMATICAL STATISTICS

MU ALPHA THETA

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, San Diego, California, April 12–15, 1978.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Peachtree Plaza Hotel, Atlanta, November 7–9, 1977.

PI MU EPSILON

SCHOOL SCIENCE AND MATHEMATICS ASSOCIATION, William Penn Hotel, Pittsburgh, Pennsylvania, November 10–12, 1977.

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, Hilton Inn, Albuquerque, New Mexico, October 31–November 2, 1977.

Prindle, Weber & Schmidt Undergraduate Mathematics Competition I

Announcement of Winning Entries

Prindle, Weber & Schmidt is pleased to announce the winners of the Undergraduate Mathematics Competition I:

Rose-Hulman Institute of Technology, Terre Haute, Indiana

Richard K. Wolf *Problem solved:* #4

Western Maryland College, Westminster, Maryland

Jeffrey A. Gates, Sherri Sheckler *Problems solved:* #1, #4, #5

University of Southern California, Los Angeles, California

Niles D. Ritter *Problems solved:* #1, #3, #4, #5

Pennsylvania State University, University Park, Pennsylvania

Charles A. Morin, Peter Pappas, J. Eric Brosius *Problems solved:* #1, #2, #3, #4, #5

University of Pittsburgh, Pennsylvania
Math 101, Problem Seminar in Mathematics

Dan Fitzsimmons, David Henry, Marie Lesley, Gary Horneman, Jeff Rinscheid
Problems solved: #1, #4,

Southern Massachusetts University, North Dartmouth, Massachusetts
Math 487, Section 1870

Ann Marie Ameen, Joseph Barckett, Joann Bates, Theresa Belli, Francis Bonaparte, Deborah Catlow, Brian Faria, Richard Farias, Michelle Gamache, Kathleen Grace, Louisa Lamaire, Michael Meredith, Susan Pereira, Richard Rousseau, Thomas Spaulding, Monica Ventura, Sandra Viveiros, David Surprenant
Problems solved: #1, #2, #3, #4, #5

Westminster College, New Wilmington, Pennsylvania

Larry Henderson, Brian Jones *Problems solved:* #1, #3, #4

We regret the delay in the mailing of the competition and the announcement of the winners.

Look for the PWS Undergraduate Mathematics Competition II on the back covers of the January 1978 issues of the American Mathematical Monthly and the Two-Year College Mathematics Journal. In addition, the new competition will appear in most of our direct mail brochures.

We wish to thank all those who participated.

We express our special appreciation to Dr. Howard W. Eves and the Mathematics Department of the University of Maine, Machias, Maine, for their enthusiastic assistance.

THE MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.
Washington, DC 20036

THE AMERICAN
MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

AN OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 84 NUMBER 9

CODEN: AMMYAE

CONTENTS

(Continued on inside cover)

NOVEMBER 1977

NOTICE TO AUTHORS

Research papers per se are unsuitable; see statement of policy (Vol. 81, p. 1).

Please follow the format in current issues of the Monthly. Manuscripts must be legibly typewritten or reproduced from typewritten copy, double spaced with wide margins, and on one side of the paper. Three copies should be submitted to the appropriate editor and one kept by the author as protection against loss. The author's full address *must* appear at the end of the manuscript.

Backlog: Main Articles 19 months, Math. Notes 18 months, Research Problems 11 months, Classroom Notes 18 months, Math. Education 17 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to R. P. Boas, Department of Mathematics, Northwestern University, Evanston, IL 60201; NOTES, etc.: to the corresponding Associate Editor; REPRINT PERMISSION: to Leonard Gillman, Mathematical Association of America, University of Texas, Austin, Texas 78712 (see also the copyright notice below); ADVERTISING CORRESPONDENCE: to Raoul Hailpern, Mathematical Association of America, SUNY at Buffalo, Buffalo, N.Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. Willcox, Mathematical Association of America, 1225 Connecticut Ave., N. W., Washington, D.C. 20036; BACK ISSUES: Contact P. and H. Bliss Co., Middletown, CT 06457.

Annual dues for members of the Association, including a subscription to the American Mathematical Monthly, are $18.00 for each of the first two years of membership and $22.00 thereafter. Student Membership is available with annual dues of $12.00. For nonmembers the subscription price is $25.00.

Published by the Association at Washington, D.C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D.C., and additional mailing offices.

PRINTED IN THE UNITED STATES OF AMERICA

POSTER SESSIONS AT THE JANUARY 1978 MEETING OF THE ASSOCIATION

At its meeting at Atlanta, Georgia, January 6–8, 1978, the Association will sponsor a poster session in addition to the customary sessions of invited talks. The poster session is for contributed papers. Each contributor prepares a visual display. The displays are posted during an early part of the meeting and, at a later session, the author is present for discussions with interested persons.

Contributions from all areas of mathematics education at the collegiate level are invited for this poster session. Two-year college mathematicians are especially invited to contribute papers to this poster session.

Abstract forms and additional information are available upon request from David Roselle, Secretary, Mathematical Association of America, VPI & SU, Blacksburg, VA, 24061. The deadline for completed abstracts is December 10, 1977.

APPLICATIONS OF CONFORMAL MAPPING TO POTENTIAL THEORY THROUGH COMPUTER GRAPHICS

DONALD T. PIELE, MORRIS W. FIREBAUGH AND ROBERT MANULIK

Abstract. Many problems in the physical sciences can be described by potentials and thus can often be solved through the application of conformal mapping. The primary limitation of this approach, especially in the classroom, is the tedious calculation and plotting that a visual image entails. We have eliminated this difficulty by the use of interactive computer graphics. By observing directly the direction and speed of mappings as they occur, students develop a deeper understanding of the nature of complex mappings which, without the computer, only experts achieve. We believe that this is just one example of many problems in physics and mathematics in which the computer can serve as an "analytical microscope" of exceptional power, clarity, and resolution.

1. Introduction. Potential theory is used in the mathematical investigation of a great variety of natural phenomena. It is particularly valuable in describing fluid dynamics, heat flow, and electromagnetic and gravitational fields. If the physical geometry of the problem can be reduced to two dimensions, the full power of complex analysis may be brought to bear, and many otherwise difficult or intractable problems readily yield a solution.

Nearly all upper level college textbooks in complex analysis for students in mathematics and the physical sciences contain a chapter on the applications of conformal mapping to the problems in potential theory. Here students are exposed, possibly for the first time, to one of the more elegant examples in mathematics where the laws of nature and new mathematical concepts are combined to make a "whole" greater than the sum of its parts. Historically, this occurred very naturally. Lagrange (1736–1813) first isolated the potential function in his studies of fluid flow. Riemann (1826–1866) developed many of his ideas about complex functions through his studies of the flow of electrical currents along a plane [1].

To appreciate this synthesis, the student needs more than an analytical understanding alone. Unless visualized, conformal maps are difficult to fully understand. But even the simplest mappings are hard to draw and we are discouraged from developing our understanding which so often is pictorially dependent. For example, we all have a mental picture of the graph of $\sin(x)$ for x real, because we have constructed it many times. But how many can visualize the graph of $\sin(z)$ for a complex z on a pair of grid lines $x = c$ and $y = c$? We all have seen, at some time, the argument proving that conformal maps conserve the magnitude and sense of angles, but how many times have we observed it other than in textbooks? Even more regrettably, few of the visually striking flow patterns or equipotential graphs that appear in texts are ever explained in terms of how they were constructed. Occasionally, an equation is given, but its implicit form makes it difficult to handle and students are never seriously asked to construct the level curves with it.

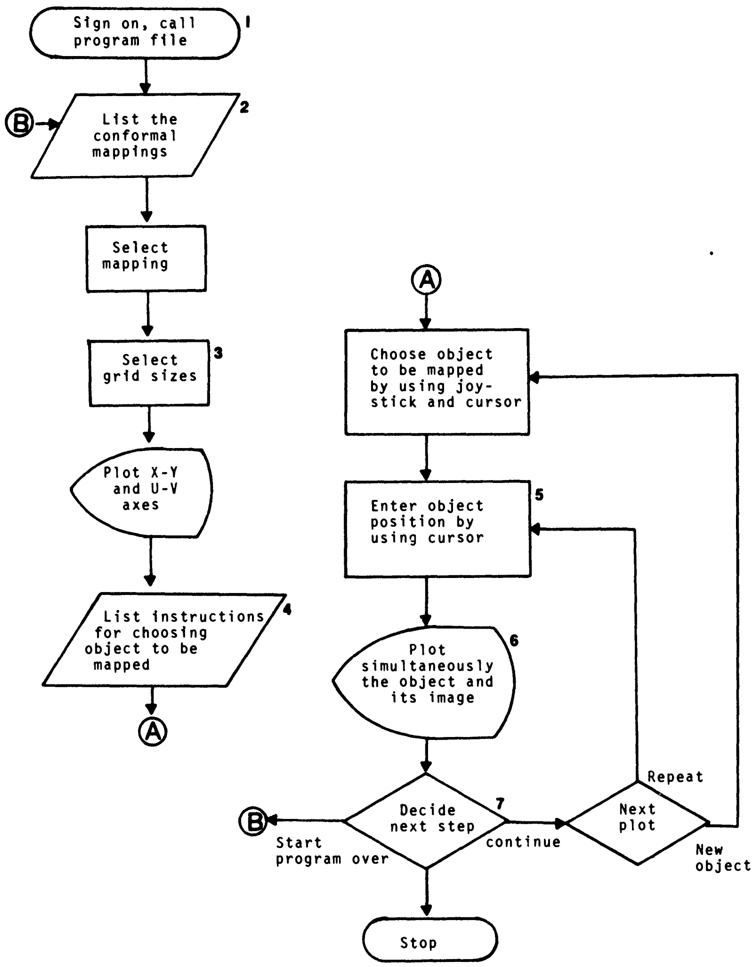


FIG. 1. Operational flowchart for conformal mapping program.

In this paper we shall show how interactive computer graphics can be used to view conformal transformations. We shall apply this technology to construct flow lines and equipotential lines for potentials associated with problems in fluid dynamics and electrostatics. We conclude by reconstructing a proof of Steiner's theorem using conformal mapping.

Without the graphical capabilities of computers, our approach, while still possible, is laborious, impractical and limited by the static nature of the resulting map. The dynamic capabilities of computer graphics can be used to simultaneously convey information about line identity, stretching, and the sense of mappings which, in textbooks, is difficult to present.

2. Computer graphics system and program.

GRAPHICS HARDWARE. We begin with a brief description of the graphics system on which our program was developed, since it differs in several respects from other instructional computer graphics projects [2, 3].

The graphics terminal is a Princeton Electronics Products 801 video terminal [4] with a 10″ × 10″ screen with 1000 line resolution. The terminal is teletype compatible and stores information locally on a LITHOCON silicon storage tube. Such memory permits selective erase of single characters,

character lines, points, and vectors. The terminal has a joystick driven cursor for entering position data directly from the screen, and hardware zoom for $4X$, $9X$, and $16X$ magnification. The most significant hardware feature, from the instructional point of view, is the video compatibility of the system. The graphics information displayed on the screen may be simultaneously presented on closed circuit television, projected by video projector, or recorded on video tape. With minimal programming effort, the video graphics output may be transmitted to a plotter for high resolution hardcopy graphics output, **[5]**.

Our terminal is linked by telephone to the UNIVAC 1110 at the Madison Academic Computing Center. Programs are available as public files to all users throughout the University of Wisconsin system.

CONFORMAL MAPPING PROGRAM. *Operational Characteristics.*

In Fig. 1 we present the operational flowchart for the conformal mapping program. To illustrate the interactive nature of this program we will describe a typical terminal session in which an instructor is demonstrating the properties of conformal maps for a class or a student is investigating these properties for himself. The numbering below corresponds to the numbering on the flowchart.

1. The user dials the UNIVAC 1110 on the dataphone, signs on, requests the program file and executes it in four lines of instruction from the terminal keyboard.

2. The program responds with the following output text to the graphics screen.

THIS PROGRAM PERFORMS CONFORMAL MAPPING FROM THE $X-Y$ PLANE TO THE $U-V$ PLANE. IT CAN ALSO PERFORM THE INVERSE MAPPING. THE FOLLOWING IS OUR MENU OF FUNCTIONS

1) FRACTIONAL LINEAR TRANSFORMATION
2) EXP(Z)
3) LN(Z)
4) COSINE(Z)
5) SINE(Z)
6) $W = \text{EXP}(Z) + Z$
7) HYPERBOLIC SINE(Z)
8) HYPERBOLIC COSINE(Z)
9) $W = Z**2$
10) $W = Z + 1/Z$ (JOUKOWSKI TRANSFORMATION)
111) $W = i*\ (\text{EXP}(Z)+1)/(\text{EXP}(Z)-1)$
12) CURRENTS IN A HOUSE WALL
13) PARALLEL PLATES, SOURCE AT 1, SINK AT INFINITY
14) FLOW IN A REACTOR (SOURCE AND SINK)
15) LINE CHARGE BETWEEN PARALLEL PLATES
16) FLOW AGAINST A WALL
17) FLOW IN A REACTOR
18) FLOW AGAINST A BARRIER
19) FLOW AGAINST A POLE

TYPE THE NUMBER OF THE FUNCTION YOU WISH TO USE

The user responds with the desired transformation number.

3. The program next requests the grid size to be used in plotting the $X-Y$ and $U-V$ axes. These are supplied by the user, depending on the function requested. If the optimum range is not obvious from inspection of the function, it is quickly learned by trial and error. Upon input of the grid size, the $X-Y$ and $U-V$ axes are immediately plotted on the top two-thirds of the screen.

4. On the bottom one-third of the screen the following instructions appear:

CHOOSE OBJECT TO BE MAPPED BY PUTTING CURSOR IN BOX AND PUSHING "END" BUTTON

☐ POINTS: GIVE ONE POINT BY POSITIONING CURSOR, PUSH END.

☐ VERTICAL LINES: GIVE ONE POINT ON LINE.

☐ HORIZONTAL LINES: GIVE ONE POINT ON LINE.
☐ ARBITRARY LINES: GIVE TWO POINTS ON LINE.
☐ CIRCLES: GIVE CENTER POINT, THEN ANY POINT ON CIRCLE.
☐ LINE SEGMENTS: GIVE TWO END POINTS.
TO START PROGRAM OVER OR QUIT, PUSH "HOME", THEN "END".

The user positions the cursor to select an object by manipulating the joystick and pressing a button.

5. The points which locate the object are next entered by using the joystick to position the cursor at the appropriate points on the $X-Y$ plane (or $U-V$ plane if the inverse is desired).

6. The computer plots simultaneously the object, which the user defined, on the $X-Y$ plane and its image under the conformal mapping-on the $U-V$ plane. Steps 5 and 6 are repeated as often as required to display the desired pattern. As an example, we show in Fig. 2 a time lapse series of pictures of the display as horizontal lines on the $X-Y$ plane are mapped under the transformation $W = Z + e^Z$.

7. After mapping each object, the user has the option of mapping a different object, starting the program over using a different transformation, or terminating the run.

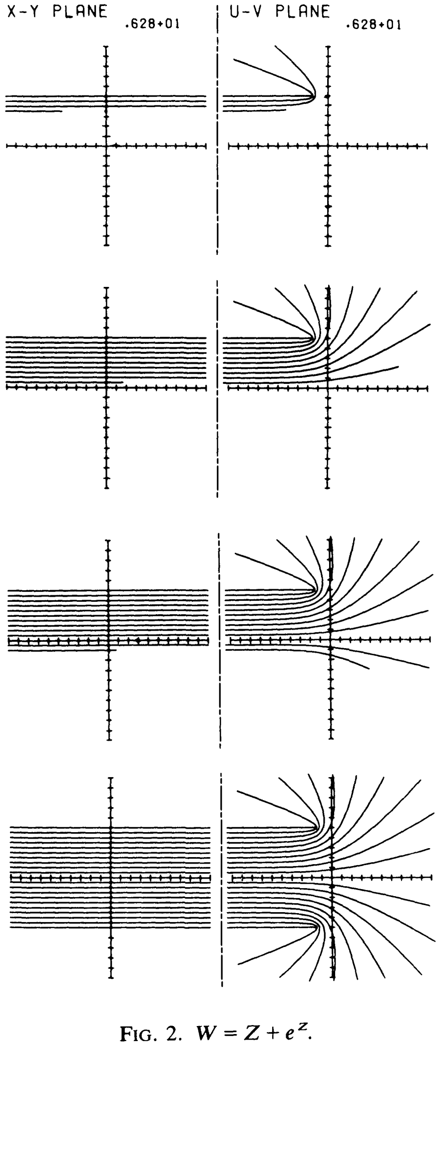


FIG. 2. $W = Z + e^Z$.

Program Details. Our program was written in FORTRAN and made extensive use of graphics subroutines available on the UNIVAC 1110. A similar program could be written in BASIC using a minimal amount of graphics subroutines since the graphics capabilities required to write our program are elementary; these include the ability to read and write text and the ability to read the cursor position. The general structure of the program is apparent from the operational flowchart shown in Fig. 1. However, this flowchart illustrates only the general outline of the program. We now examine in more detail programming problems that must be solved in order to construct satisfactory mappings.

In any graphics program written for displaying functions or mappings one encounters the problem of determining the proper scaling for both the domain and range axis. In our program, where we allow the user to choose different dimensions for the x–y plane and u–v plane, algorithms for scaling are especially important. This problem can be handled with graphics system subroutines, if they are available, or by writing algorithms directly into the program. This is not difficult to do since all problems of positioning and scaling a coordinate system can be reduced to finding the appropriate two-dimensional linear transformation.

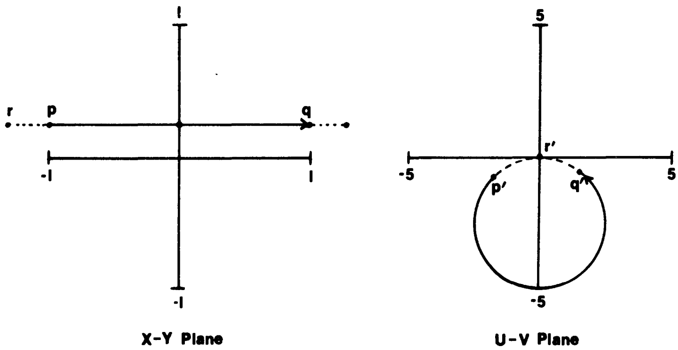


FIG. 3. The mapping $W = 1/Z$ which maps the point at infinity to zero.

Another problem involved in constructing conformal mappings is best illustrated by an example. In the fractional linear transformation $w = 1/z$ an entire line in the x–y plane is transformed into a circle in the u–v plane with the point at infinity being mapped to zero. This presents the problem of determining where to start and stop the sweep of the line in the x–y plane and what step size to use so that the image in the u–v plane is both complete and smooth. In Figure 3, if we start the sweep of a horizontal line in the x–y plane at $p = (-1, .2)$ and end it at $q = (1, .2)$ then the image under the mapping $w = 1/z$ is that portion of the circle from p' to q' shown in the u–v plane. To make an entire circle in the range it is necessary to either change the dimensions of the x–y axis or begin the map out of view to the left of $(-1, .2)$. In most cases the later is preferred because we keep the flexibility of choosing the dimensions we want for the domain of each function. To find the starting point in this case, a series of guesses are made and the image of each guess computed. The image of the point at infinity is also calculated and compared with the image of each guess. When the two are sufficiently close together the last guess is used as the starting point. For example, a starting point of $r = (-20, .2)$ puts the image r' within approximately 1/40 of zero. The line segment from r to p does not appear on the screen, but the image from r' to p' does; similarly for q' to r'.

The last problem is determining what step size to use in the sweep of the horizontal line. In the mapping $w = 1/z$ it would be impractical to use a uniform step size because the image near zero would move very slowly and the image near $(0, -5)$ would move rapidly through large jumps. To achieve a smooth circle drawn at a uniform rate, a step size is used that decreases exponentially as we approach the y-axis and increases exponentially as we move away. Once the start and end points are found, the program steps through the object in the x–y plane. Image points in the range and line segments

between them are drawn following the construction of each line segment in the domain. Thus, each object and its image appear simultaneously on the screen.

3. Applications of conformal mapping to potential theory. We now examine how our program was applied in two major areas of physics, fluid dynamics and electrostatics, and conclude with a theorem in mathematics.

FLUID DYNAMICS. The value of complex functions in the study of potentials is derived from the fact that if

$$w(z) = w(x, y) = u(x, y) + iv(x, y)$$

is an analytic function of the complex variable $z = x + iy$, then both u and v are potential functions and satisfy Laplace's equation. In addition, if $\phi(x, y)$ is a potential function then a conjugate potential $\psi(x, y)$ exists such that

$$F(z) = \phi(x, y) + i\psi(x, y) \tag{1}$$

is an analytic function. The function $F(z)$ is called the *complex potential* associated with the real potential ϕ, and we shall make repeated use of it in this paper.

Let us review the role that real and complex potentials play in Newtonian or ideal fluid flow. If fluid motion is irrotational and incompressible then its velocity field is determined by a velocity potential ϕ [**6**]. The associated complex potential (1) holds the key to a complete description of the flow field. Specifically, the conjugate of its first derivative, $\overline{F'(z)}$, is the velocity field; the streamlines, which trace the movement of the fluid in the field, are the level curves $\psi(x, y) = c$; and the level curves $\phi(x, y) = c$ are perpendicular to the streamlines and link points together with the same velocity (isovels).

The simplest type of fluid flow is uniform and horizontal and is characterized by the complex potential $F(z) = z$. Here the conjugate potential $\psi(x, y) = y$, and the streamlines are the horizontal lines $y = c$. The speed of the flow is uniform, $|F'(z)| = 1$.

To study more elaborate fluid flows, we again call upon complex functions. It is a routine exercise in complex analysis [**7**] to show that a conformal mapping (i.e., a complex analytic function $w(z)$ with a nonvanishing derivative) transforms velocity potentials and streamlines in the z plane into velocity potentials and streamlines in the w plane. Thus, flows that are intuitively understood in simple regions can be used to study flows in more complicated regions which are conformally equivalent.

Flow around a cylinder: We apply this idea to a well-known example in fluid dynamics—flow around a cylinder placed perpendicular to a uniform flow field. The region R_w outside the unit circle is mapped conformally onto the region R_z, outside the slit $[-2, 2]$ by the Joukowski Transformation.

$$z = 1/w + w. \tag{2}$$

The potential for uniform horizontal flow in the z plane, $F(z) = z$, corresponds to the potential $G(w) = w + 1/w$ for flow around the circle in the w plane. To construct the streamlines for G, we invert the Joukowski transformation (2) and solve for w

$$w = \tfrac{1}{2}(z + (z^2 - 4)^{\frac{1}{2}}). \tag{3}$$

A single valued branch for (3) can be defined [**8**] for the region R_z. We are now ready to let the computer take over and construct the streamlines for the potential $G(w)$. As z traverses a flow pattern in R_z, the image point w, determined by function (3), traverses the corresponding flow pattern in R_w. This is illustrated in Fig. 4 by a series of pictures taken from the computer graphics screen.

An explicit representation of the level curves can be found in polar form by substituting $w = re^{i\theta}$ in (2) and setting the imaginary part equal to a constant. For the Joukowski potential we find [**9**]

$$\psi(r, \theta) = (r - 1/r)\sin\theta = c. \tag{4}$$

However, even in this relatively simple example, the implicit representation (4) is not very convenient for actually constructing the flow lines. For more complicated potentials it becomes even less useful.

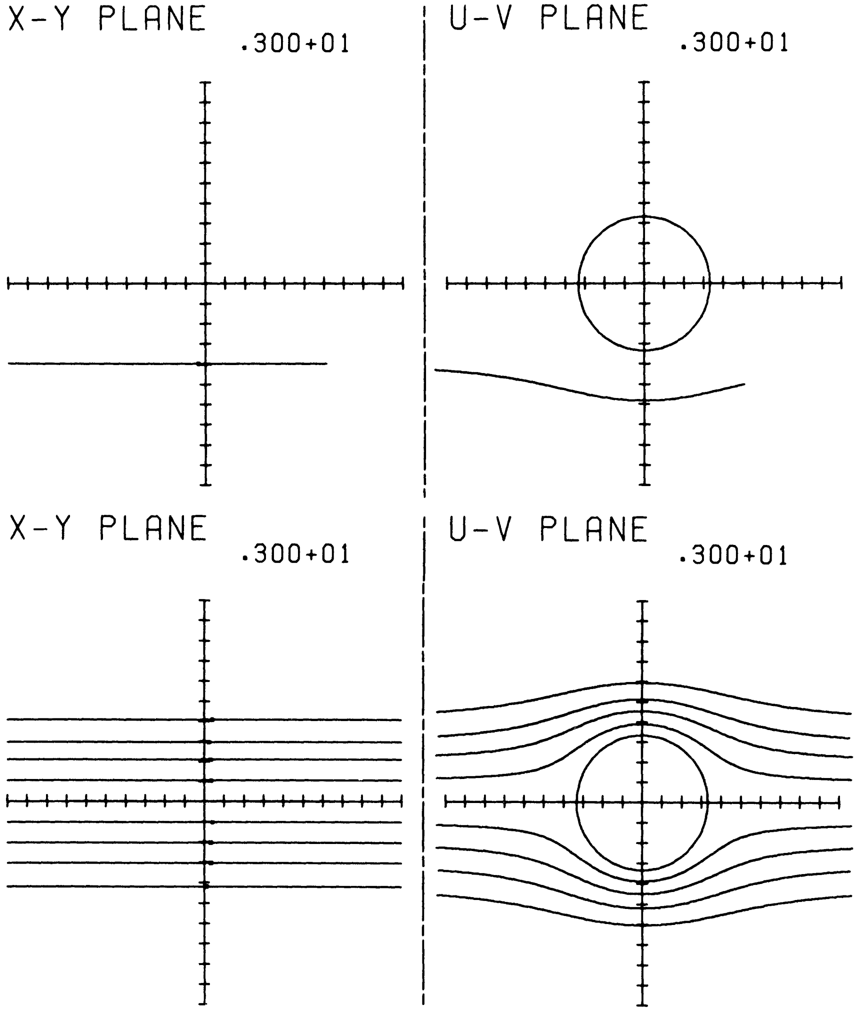


FIG. 4. Flow around a cylinder constructed with transformation (3).

The Airfoil: We carry the Joukowski transformation one step further and study the pure irrotational flow around an airfoil. The unit circle is mapped onto the line segment $[-2,2]$ in the z plane by (2). If a circle is centered near the origin and passes through the point $(1,0)$ then its image under (2) resembles the boundary of an airfoil as shown in Fig. 5. The shape of the airfoil can be changed by moving the center of the circle. To construct the streamlines for flow over the airfoil, we simply transform the streamlines around the cylinder in Fig. 4 into the corresponding streamlines for the airfoil under the Joukowski transformation

$$z' = w' + 1/w'. \tag{5}$$

The entire job can be done with one conformal mapping

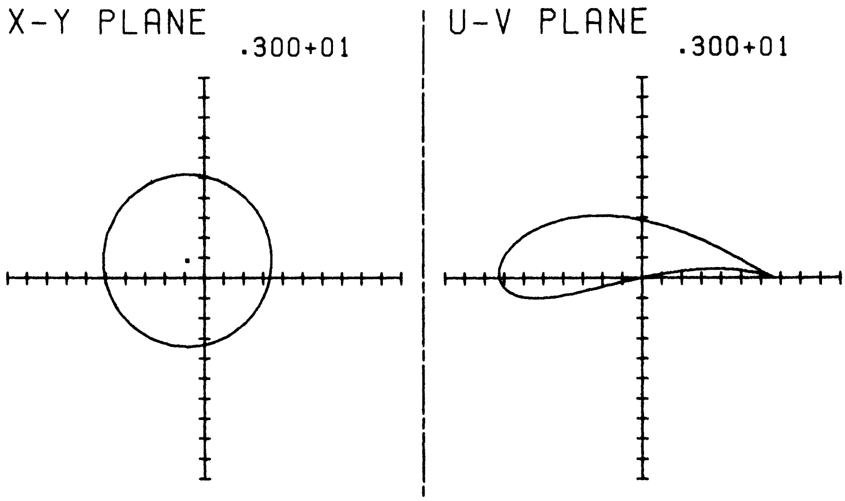


FIG. 5. The airfoil constructed with the Joukowski Transformation (2).

(6) $$w' = \frac{r(z + (z^2 - 4)^{\frac{1}{2}}) + 2z_0'}{2} + \frac{2}{r(z + (z^2 - 4)^{\frac{1}{2}}) + 2z_0'}$$

that transforms the region R_z outside of the slit $[-2, 2]$ onto the region $R_{w'}$ outside of the airfoil. This mapping is easily constructed by composing the conformal mappings (2), (5) with the translation and contraction of the circle centered at z_0' onto the unit circle centered at the origin given by

(7) $$w = (z' - z_0')/r.$$

Now the computer graphics program takes over. By traversing the horizontal lines in the complex z-plane, the image points in the w'-plane given by (6) map out the streamlines of the flow over the airfoil in a uniform horizontal flow as shown in Fig. 6. With the computer graphics capabilities, the airfoil can be easily studied in a variety of flow fields. Fig. 7(a) shows the streamlines for uniform flow against the airfoil at the incidence angle α. The Joukowski transformation can also be used to solve a related problem. By choosing $z_0' = 0$, the airfoil degenerates into the line segment $[-2, 2]$ so we can generate the flow lines around a flat plate shown in Fig. 7(b) using the same mapping (6).

Note that to more closely simulate the actual flow over an airfoil, it is necessary to introduce circulation [**10**].

Flow out of an open channel: Our final example in fluid dynamics illustrates the possibility of constructing streamlines for a flow problem using conformal mapping and computer graphics even when an explicit representation for the complex potential is impossible. Consider the region R_z between the lines $y = \pi$ and $y = -\pi$ in the complex z plane. This unbounded region is mapped conformally [**11**] by

(8) $$w = z + e^z$$

onto the w plane minus the two half lines $-\infty < u \leqq -1, v = \pi$ and $-\infty < u \leqq -1, v = -\pi$. This region is denoted by R_w. Ideal fluid flow in the infinite channel R_z is of course horizontal. The flow lines through the open channel R_w corresponds to these horizontal lines under the conformal mapping (8) as shown in Fig. 8(a). To find the complex potential for this flow, equation (8) would have to be inverted and z written as a function of w. Such an explicit inversion is impossible using elementary functions.

Finally, we observe that the flow pattern for an open channel can also be interpreted as the

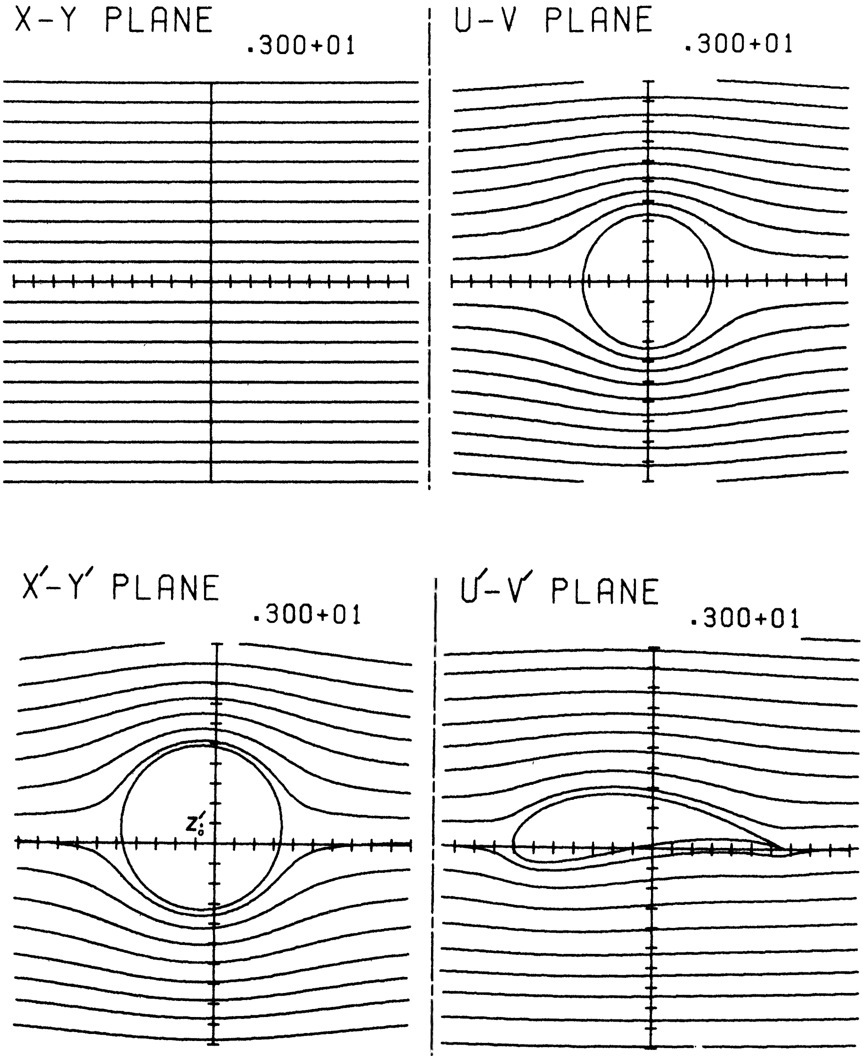


FIG. 6. Irrotational flow over the airfoil constructed with transformation (6).

equipotential lines formed at the edge of a parallel plate capacitor. In this electrostatic setting, the conjugate potential may be used to construct the electric field lines as shown in Fig. 8(b). This leads us naturally to our next topic.

ELECTROSTATICS. Through the pioneering work of Poisson and Gauss, the potential function was used by the beginning of the nineteenth century to describe fields in electrostatics and magnetostatics. However, very little was known about potential functions in general until George Green (1793–1841) undertook a thorough study of their mathematical properties [**12**]. In his privately published book *An Essay on the Application of Mathematical Analysis to the Theories of Electricity and Magnetism*, (1828), Green introduced a special function—later named the Green's function by Riemann—from which he inferred the possibility of solving a wide class of boundary value problems, known as the Dirichlet problem. We mention the Green's function here because it occurs naturally in many problems in electrostatics. In fact, Green inferred the existence of his special function from the Faraday cage experiment where static charge is always induced on a closed, grounded, conducting

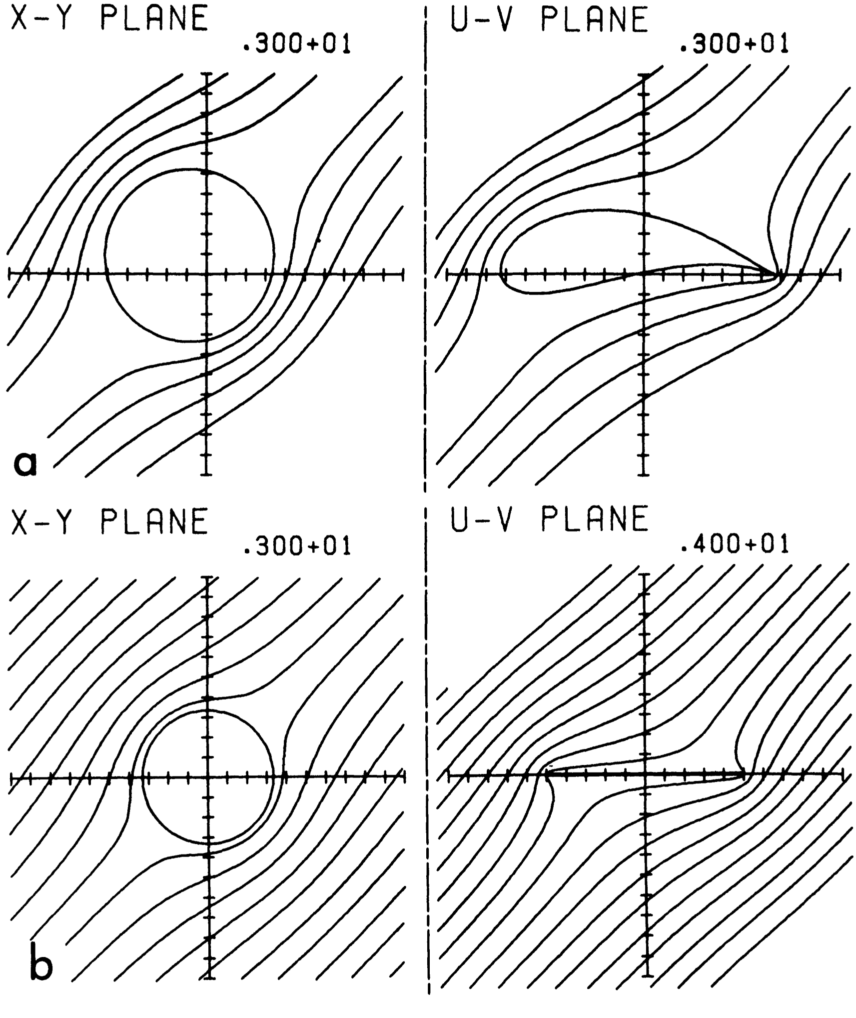


FIG. 7a. Flow against an airfoil at the incidence angle $\alpha = 45°$.

FIG. 7b. Flow against a flat plate at the incidence angle $\alpha = 45°$.

surface by a point charge within the conductor, such that the resulting combined potential vanishes on the surface [**13**].

We consider electrostatic fields for which complex analysis can be successfully applied. If the field $\vec{E}$ has a zero component in the z-direction (i.e., cylindrical geometry), then it is sufficient to study the field in a plane cross section where we can use the complex notation. In the charge-free region away from an electrode or conducting surface, the field $\vec{E}$ is determined by a real potential function $\phi(x, y)$ through the gradient

(9) $$\vec{E} = -\left(\frac{\partial\phi}{\partial x} + i\frac{\partial\phi}{\partial y}\right).$$

Since ϕ is harmonic in the charge-free region, we can construct the complex potential [**14**]

(10) $$F(z) = \phi(z) + i\psi(z)$$

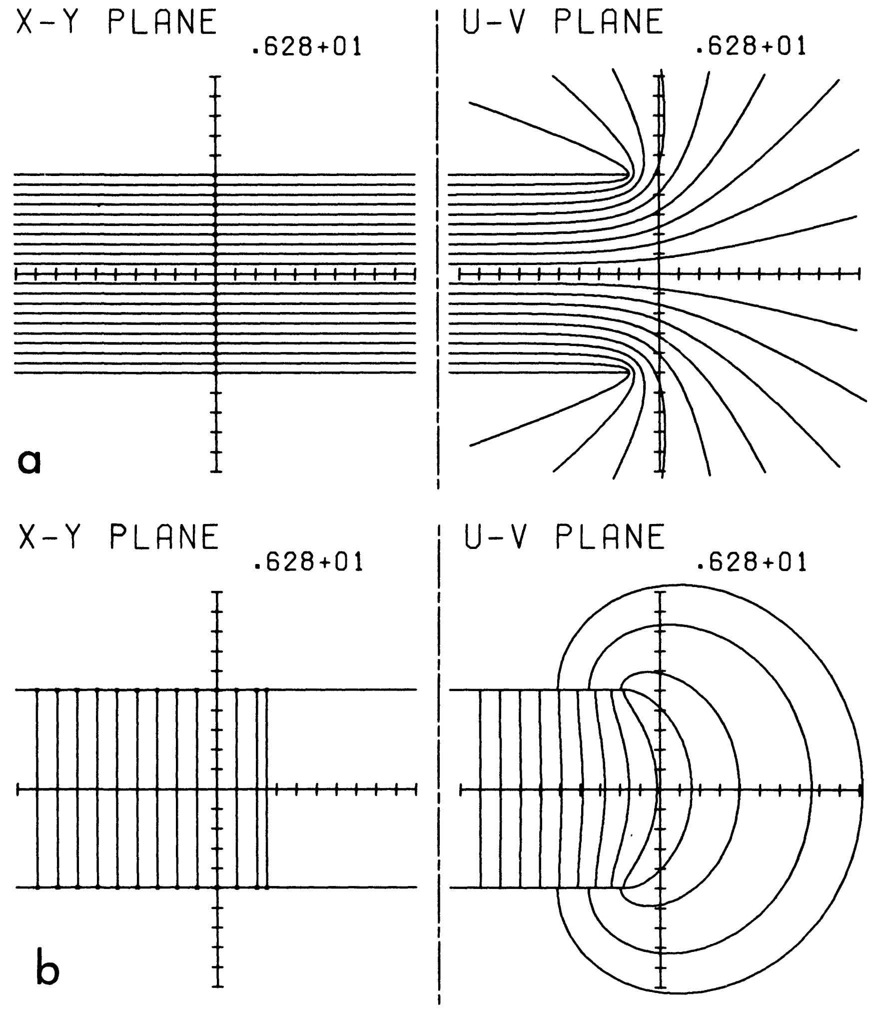


FIG. 8a. Flow out of an open channel or the equipotential lines formed at the edge of a parallel plate capacitor.

FIG. 8b. Lines of force at the edge of a parallel plate capacitor. Both a and b were constructed with transformation (8).

just as we did in fluid dynamics. Again we find that the potential $F(z)$ supports many properties of the field $\vec{E}$. For example, $\vec{E} = F'(z)$; the level curves $\phi(z) = c$ are the equipotential lines for $\vec{E}$; and the level curves $\psi(z) = c$ are the lines of force for a unit charge in the field $\vec{E}$.

To construct these level curves using conformal mapping, we consider the complex potential $F(z)$ as a transformation from the z-plane to the w-plane and write

$$w = F(z) = \phi(z) + i\psi(z). \tag{11}$$

If we express w in terms of its components, $w = u + iv$, the level curves $\phi(z) = c$ in the z-plane correspond to the vertical lines $u = c$ in the w-plane, and the streamlines $\psi(z) = c$ in the z-plane corresponds to the horizontal lines $v = c$ in the w-plane. Thus if we invert the transformation (9) and write z as a function of w

$$z = F^{-1}(w), \tag{12}$$

the images of the horizontal and vertical lines in the w-plane are precisely the streamlines and equipotential lines for the field $\vec{E}$ in the z-plane.

Line Charges: Consider the electrostatic field generated by two line charges of opposite sign running perpendicular to the x–y plane and passing through the points z_0 and $\bar{z}_0$. Clearly, the field $\vec{E}$ has a zero component in the z-direction. The two-dimensional real potential is given by **[15]**

$$\phi(z) = \frac{1}{2\pi} \log \frac{|z - z_0|}{|z - \bar{z}_0|}, \tag{13}$$

where $|z - z_0|$ denotes the distance between the point z and the line charge at z_0, and $\bar{z}_0$ is the complex conjugate of z_0.

The associated complex potential

$$F(z) = \frac{1}{2\pi} \log \frac{z - z_0}{z - \bar{z}_0} \tag{14}$$

follows immediately from the definition of the complex logarithm,

$$\log w = \log |w| + i \arg w. \tag{15}$$

Also, the conjugate function $\psi(z)$, can be derived from (15)

$$\psi(z) = \arg \frac{z - z_0}{z - \bar{z}_0}. \tag{16}$$

We are now ready to invert the complex potential (14) and solve for z in terms of w

$$z = \frac{\bar{z}_0 e^{2\pi w} - z_0}{e^{2\pi w} - 1}. \tag{17}$$

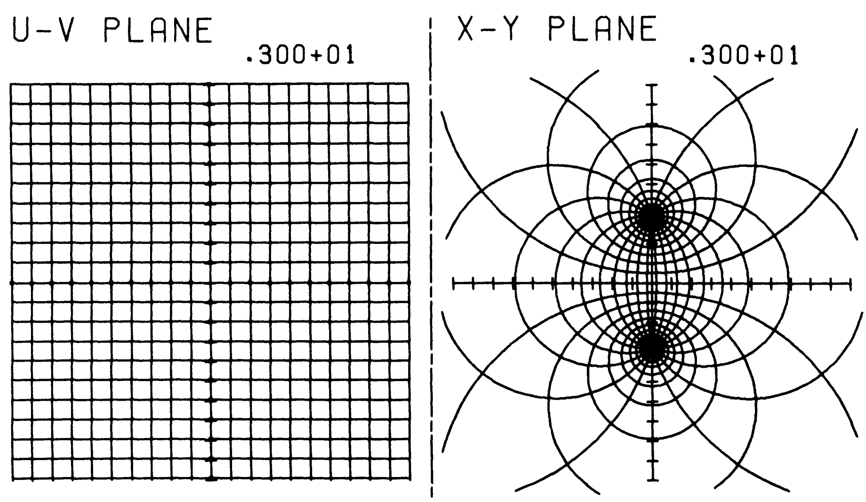


FIG. 9. The equipotential lines and the lines of force for two line charges of opposite sign, constructed with transformation (17).

When we traverse the vertical and horizontal lines in the w-plane, the corresponding equipotential lines and force lines for the potential ϕ are constructed in the z-plane. This is a simple task for our graphics program and is illustrated for $z_0 = i$ in Fig. 9. Several remarks are in order.

1. The horizontal line, $v = 0$, in the w-plane is mapped onto the x-axis in the z-plane. Thus ϕ does vanish on the x-axis.

2. As a function of z and z_0, ϕ is the Green's function for the Dirichlet problem in the upper-half plane.

3. The potential ϕ has a meaning in fluid dynamics. It corresponds to the potential due to a source at z_0 and a sink at $\bar{z}_0$.

4. The potential ϕ has another electrostatic interpretation. It is the potential due to a line charge at z_0, in the presence of a grounded "infinite" conducting plate on the x-axis.

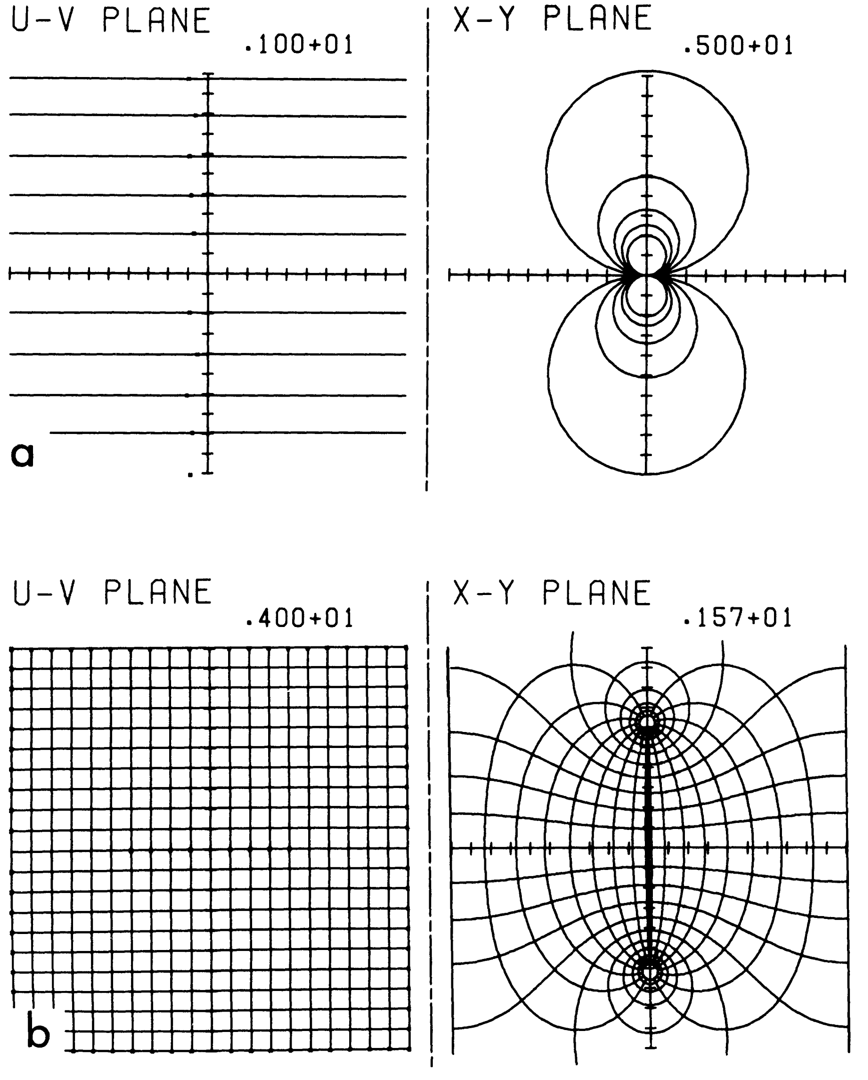


FIG. 10a. The dipole constructed with transformation (18).

FIG. 10b. Currents in a house-wall constructed with transformation (24).

The Dipole: If we allow z_0 to approach the origin in the above example and at the same time adjust the charges so that the dipole moment remains constant, then we can derive [**16**] the complex potential for the dipole

$$F(z) = 1/z. \tag{18}$$

This is an especially easy function to invert and construct the equipotential lines as shown in Fig. 10(a).

Currents in a house-wall: As a final example in electricity, we shall construct the equipotential and current flow lines in an idealized conducting house-wall containing a positive and negative line source in the interior of the wall [17]. Again we assume the wall is infinitely high so we can reduce the problem to a two-dimensional one. This problem is similar to the line charge and conductor problem illustrated in Fig. 9 with the added condition that all flow lines are confined to an infinite slab parallel to the y axis. Mathematically this means that the normal derivative of the potential must vanish at the surface of the walls. This condition is achieved through an application of the method of images which places unit charges periodically along two lines parallel to the x-axis. Assuming a wall width of π and line charges at i and $-i$, the complex potential in the house wall becomes

$$F(z)=\log\frac{\sin(z-i)}{\sin(z+i)}. \tag{19}$$

To construct the equipotential lines and the lines of flow, we must once again invert the potential. First translate (19) into exponential form

$$e^{w}=\frac{\sin(z-i)}{\sin(z+i)}, \tag{20}$$

where $w=F(z)$. Next, use the trigonometric identity

$$\sin(z\pm i)=\sin(z)\cos(i)\pm\cos(z)\sin(i) \tag{21}$$

to expand and rewrite (20). Finally, collect like terms and divide through by $\cos(z)$ to get

$$\tan z=\tan(i)\cdot\frac{1+e^{w}}{1-e^{w}} \tag{22}$$

which is easily inverted to

$$z=\arctan\left(\tan(i)\cdot\frac{1+e^{w}}{1-e^{w}}\right). \tag{23}$$

This function can be expressed with complex logarithms as

$$z=1.32\log\tfrac{1}{2}\left(e^{w}+(e^{2w}+4i)^{\frac{1}{2}}\right). \tag{24}$$

Using equation (24) in our graphics program, we can easily construct the level curves for the potential inside a house-wall as shown in Fig. 10(b).

Again the potential that we have investigated is a Green's function. Specifically, the real potential

$$\phi(z,i)=\frac{1}{2\pi}\log\left|\frac{\sin(z-i)}{\sin(z+i)}\right| \tag{25}$$

is the Green's function at $z_0=i$ for the infinite well $-\pi/2\leqq x\leqq\pi/2, y\geqq 0$ under the mixed Dirichlet and Neumann boundary conditions

$$\phi=0 \text{ on the } x\text{-axis, } y=0,\ \partial\phi/\partial n=0 \text{ on the walls } x=\pi/2 \text{ and } x=-\pi/2. \tag{26}$$

A MATHEMATICAL APPLICATION: STEINER'S THEOREM. Our emphasis in this paper has been on the applications of conformal mapping rather than on the study of conformal maps themselves. It should be obvious, however, that one can use a graphics program to develop intuition about the behavior of a large variety of complex analytic functions or illustrate properties of conformal maps. Branch cuts, which are often bewildering to the beginning student, exist throughout our investigations of inverse mappings. When we cross a branch cut for a mapping on the graphics terminal, the result is dramatic. An extraneous line appears in the image of the mapping that connects the jump discontinuity associated with every cut. Consequently, branch cuts are easy to illustrate on the graphics terminal, although occasionally a nuisance if not avoided. For example, most of the equipotential curves for the

current in a house-wall had to be constructed by two vertical line segments which carefully avoid crossing the branch cut on the x-axis.

For the versatility necessary to study conformal maps, we have built into our program the ability to generate the images of more than just horizontal and vertical lines. The option of transforming circles permits us to reconstruct a proof of Steiner's Theorem.

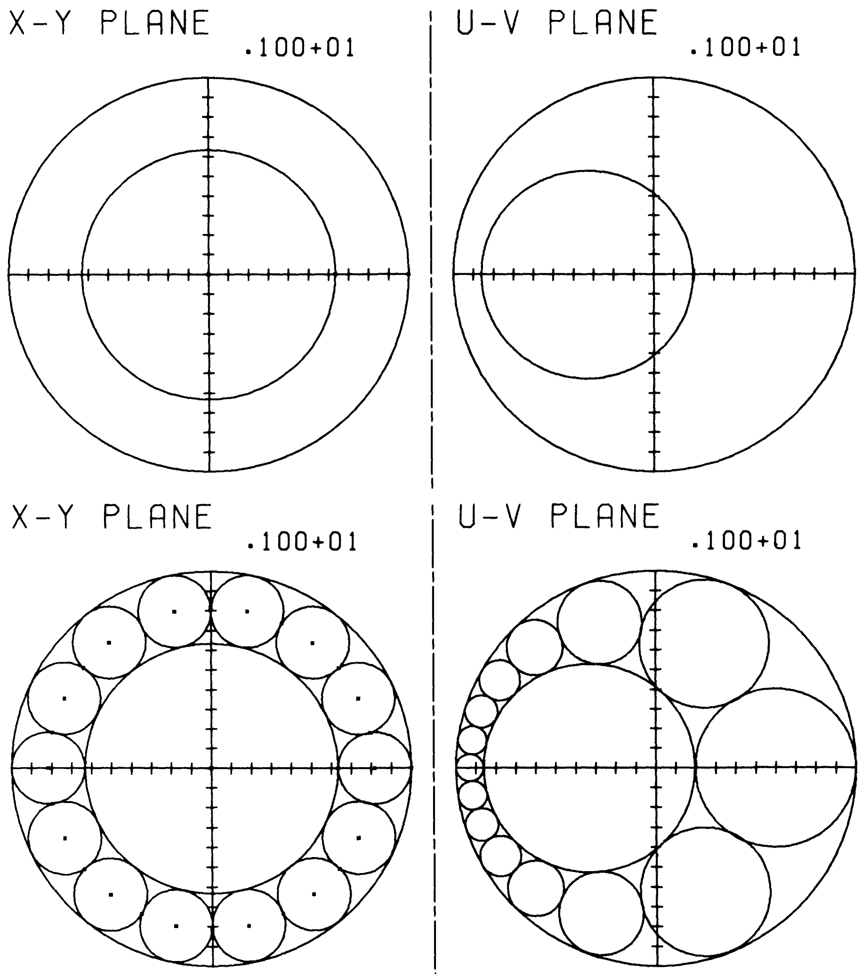


FIG. 11. The bilinear mapping (27) with $a = .5$.

Steiner's Theorem: Draw two circles in the plane, one inside the other, and draw other circles tangent to them and to one another as shown in Fig. 11. Sometimes the ring of circles fit snugly together like a set of ball bearings and the last circle is tangent to the first one. It would be natural to expect that some arrangements fit perfectly while others do not, depending upon where you place your first circle. However, Steiner's theorem contradicts our intuition and states that the ring of circles will fit perfectly for *every* position of the starting circle if it fits for *any* position of the starting circle **[18]**.

The proof of Steiner's theorem using conformal mapping can be demonstrated with our program. First, we must construct a conformal mapping that transforms the region between the two non-concentric circles onto an annular region. This is always possible **[19]** and the mapping is bilinear

$$z = \frac{a - w}{aw - 1}, \tag{27}$$

where $-1 < a < 1$. Since bilinear maps transform circles and lines into circles and lines, the ring of circles shown on the right in Fig. 11 is transformed into a ring of circles of uniform diameter inside the annulus. Of course, for the annulus, the position of the first circle has no effect on whether the circles fit snugly or not and Steiner's theorem is completely obvious. Thus, the theorem must hold for arbitrary nested circles.

Conclusion. Several educational benefits resulted from using computer graphics to construct conformal mappings. The exercise of writing the program was a significant educational experience and was done as a project in an undergraduate course on computer graphics [**20**]. Finding complex potentials associated with applied problems and computing their inverses was also a worthwhile exercise. The graphics program was used in class in an interactive mode to examine properties of individual conformal mappings. Finally, a video tape of the most interesting mappings was made for use independent of the computer and future reference [**21**].

Acknowledgment. We wish to thank Professor B. Greenebaum and C. J. Goebel for critically reading the manuscript and making several helpful suggestions.

This work was supported in part by the Parkside Faculty Research Fund.

References

1. M. Kline, Mathematical Thought From Ancient to Modern Times, Oxford University Press, New York, 1972.

2. A. Bork, The J. of Tech. Horizons in Ed. 1, 6 (1974).

3. E. Lyman, A Summary of PLATO Curriculum and Research Materials, CERL Report x-23, University of Illinois, Urbana, 1972.

4. More information on the terminal is available from Princeton Electronic Products, P.O. Box 101, North Brunswick, New Jersey 08902.

5. The figures in this paper were produced by CALCOMP plotter at the Madison Academic Computing Center.

6. R. Churchill, J. Brown, and R. Verhey, Complex Variables and Applications, McGraw-Hill, New York, 1974, 3rd ed., p. 227.

7. See Ref. 6, p. 206.

8. A. G. Sveshnikov and A. W. Tikhonov, The Theory of Functions of a Complex Variable, MIR Publishers, Moscow, 1971, p. 176.

9. See Ref. 6, p. 234.

10. R. Rothe, F. Ollendorff, and K. Pohlhausen, Theory of Functions as Applied to Engineering Problems, Dover, New York, 1961, p. 117.

11. H. Kober, Dictionary of Conformal Representations, Dover, New York, 1957, p. 117.

12. See Ref. 1, p. 682.

13. O. D. Kellogg, Foundations of Potential Theory, Dover, New York, 1953, p. 277.

14. See Ref. 6, p. 229.

15. A. G. Mackie, Boundary Value Problems, Hafner, New York, 1965, p. 167.

16. N. Levinson and R. Redheffer, Complex Variables, Holden Day, San Francisco, 1970, p. 280.

17. See Ref. 10, p. 98.

18. See Ref. 16, p. 324.

19. See Ref. 16, p. 276.

20. R. Manulik wrote the program in the undergraduate course offered by M. Firebaugh.

21. The videocassette *Application of Conformal Mapping* is available through the Media Production Center, UW-Parkside, Kenosha, Wisconsin 53141.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN-PARKSIDE, KENOSHA, WI 53141.
DEPARTMENT OF PHYSICS, UNIVERSITY OF WISCONSIN-PARKSIDE, KENOSHA, WI 53141.
DEPARTMENT OF PHYSICS, UNIVERSITY OF WISCONSIN-MADISON, MADISON, WI 53706.

FINITE SIMPLE GROUPS

JAMES F. HURLEY AND ARUNAS RUDVALIS

1. Introduction. Perhaps the most active research field in algebra today is the study of finite groups, with much of the activity centered on the study of finite simple groups. The purpose of this paper is to introduce the reader to the field of finite simple groups by indicating something of its inherent importance in group theory and, by tracing its history, to describe some of the tools which have been and are being used to attack the problem of determining all the finite simple groups. We conclude by summarizing some of the recent discoveries in the field and some of the possibilities they suggest. For those who might wish to learn more about this field we include a lengthy bibliography. However, with the volume of new material which is constantly appearing, this bibliography will soon be (if indeed it is not already) guilty of serious omissions. To keep up to date, the reader is urged to consult the *Mathematical Reviews* periodically and also the *Journal of Algebra* and *Communications in Algebra*, which contain many papers in this field.

Before beginning a description of the work that has been done on the classification problem let us first see why one is interested at all in finite simple groups, assuming that one wants to learn all one can about finite groups in general. There are a number of ways to approach this. One of the most appealing is through the analogy between finite group theory and the arithmetic of the natural number system. To bring this analogy into focus, let us consider divisibility in the set of natural numbers.

We say that the positive integer m divides the positive integer n if there is a positive integer q such that $n = mq$. In this circumstance we write $n/m = q$, and call q the quotient of n by m. Relative to this concept, which natural numbers are the most basic and simple? These are the natural numbers $p \neq 1$ whose only divisors are the trivial divisors, 1 and p. These natural numbers are called the **prime** positive integers. The central fact of elementary arithmetic theory is the theorem unsurprisingly called the Fundamental Theorem of Arithmetic, usually stated as: Every positive integer $n \neq 1$ can be expressed in the form $n = p_1^{e_1}p_2^{e_2}\cdots p_k^{e_k}$ for distinct primes $p_1, p_2, \dots, p_k$ and natural numbers $e_1, e_2, \dots, e_k$ such that each pair p_i and e_i is uniquely determined (up to possible renumbering of indices i). For purposes of building our analogy we prefer to state this in the following form, which is easily seen to be equivalent: For every positive integer $n \neq 1$ there is a sequence $n = n_0 \geqq n_1 \geqq n_2 \cdots \geqq n_{r-1} \geqq n_r = 1$ such that each n_i/n_{i+1} is a prime, and the collection of primes which so occur and their multiplicities are uniquely determined by n up to reordering. (The n_i themselves are not unique of course). Why is this theorem so fundamental? It says that the prime positive integers are not only simple in their divisibility properties, but are in fact the fundamental building blocks for the natural number system in the sense that every positive integer $n > 1$ can be obtained by multiplying together a collection of primes.

Now where is the analogy to all this in finite group theory? We recall from elementary abstract algebra that there is a group theoretic notion akin to divisibility, namely normality. For H a subgroup of G, the quotient set of G modulo H consists of the left cosets of H in G. These are precisely the equivalence classes of the relation congruence modulo H, defined by $a \equiv b \bmod H$ means $a^{-1}b \in H$. This quotient set forms a group under the natural coset multiplication $xHyH = xyH$ if and only if H is a **normal subgroup** of G (MacDonald [**33**, p. 56]), which we recall simply means $xH = Hx$ for every $x \in G$. In this case we write $H \trianglelefteq G$ or $G \trianglerighteq H$ and call the quotient set of G modulo H the **quotient group** or **factor group** of G by H, denoted by G/H. A simple example of importance occurs in the symmetric group S_n of all permutations of a set of n objects, in which the set A_n of all even permutations is a normal subgroup with $S_n/A_n \cong Z_2$, the cyclic group of two elements. Now what is the appropriate analogue of a prime in this context? A group G whose only normal subgroups are the trivial ones $\{1\}$ and G (where 1 is the identity in G). Such a group is called a **simple group**.

It is appropriate to use the term *simple* here as we can see if we approach these ideas from the

point of view of homomorphisms. In group theory as in mathematics generally, one of the keys to unlocking the nature of the objects of study lies in the structure preserving transformations of the objects. As Carroll's famous characterization of a grin suggests, the abstract essence of a mathematical object can be thought of as what remains after the accidental features of its many examples have been erased. In group theory the erasers are the group homomorphisms, mappings which transform one group G_1 into another G_2 but in the process preserve the group operation. This preservation condition, written $f(xy)=f(x)f(y)$, says that applying the operation of G_1 to x and y and then mapping the result into G_2 gives the same result as first mapping x and y into G_2 and then applying the operation of G_2 to the images $f(x)$ and $f(y)$. Throughout algebra, the kernels of homomorphisms play a prominent role, and we recall that for a group homomorphism f, Ker $f=\{x \in G_1 \mid f(x)=1 \in G_2\}$ is a normal subgroup of the group G_1, and conversely every normal subgroup N of a group G is the kernel of a homomorphism defined on G_1. The Fundamental Homomorphism Theorem for groups tells us that the homomorphic images of any group G consist precisely of the quotient groups G/K for $K \trianglelefteq G$. Referring to our example above, the two element multiplicative group $\{1, -1\}$ (which is cyclic of order 2) is the homomorphic image of the parity homomorphism f given by $f(\pi)=1$ if π is even, $f(\pi)=-1$ if π is odd. Here Ker $f=A_n$ and the isomorphism $S_n/A_n \cong \boldsymbol{Z}_2$ is immediate from the Fundamental Homomorphism Theorem. From this theorem also, we see that if G has only $\{1\}$ and G as normal subgroups, then the group G is structurally simple in the sense that its only possible homomorphic images are itself and the trivial group $\{1\}$.

Finally now, what is the analogue for finite groups of the Fundamental Theorem of Arithmetic? (At this point we really *must* restrict consideration to finite groups since the theorem does not hold for arbitrary infinite groups.) Translating the theorem in its second form above, we obtain the **Jordan–Hölder theorem for finite groups**: For every finite group G there is a sequence of subgroups $G=G_0 \trianglerighteq G_1 \trianglerighteq G_2 \trianglerighteq \cdots \trianglerighteq G_{r-2} \trianglerighteq G_{r-1} \trianglerighteq G_r=\{1\}$ such that each quotient group G_i/G_{i+1} is a simple group, and the collection of associated simple quotient groups is unique up to reordering. (Such a sequence is called a **composition series** for G.)

The reader must have noticed that while the Jordan–Hölder theorem is a classical result of finite group theory, it is not called the fundamental theorem of finite groups. Having perhaps also noticed the imperfection of our analogy in certain respects (for example, while each n_i divides n in the Fundamental Theorem of Arithmetic, each G_i need not be normal in G just because it is normal in G_{i-1}), the reader may be wondering whether it is too much to hope that the finite simple groups will also be fundamental building blocks for the system of finite groups. While it is no longer so easy to see, a group theoretic analogy with arithmetic again exists and we now proceed to describe it.

Suppose we knew all the finite simple groups. How could we use this knowledge, and the Jordan–Hölder theorem, to determine all possible finite groups? First we would write down a list $S_1, S_2, \cdots, S_r$ of simple composition factors. Then we would try to build up a tree diagram listing the possible subgroups G_i such that $G_i/G_{i+1} \cong S_{i+1}$, ending with all the possible $G=G_0$ corresponding to our given list of composition factors. The first step is easy. Since $G_r=\{1\}$, $G_{r-1}=G_{r-1}/G_r \cong S_r$. Our next step would then be to determine, knowing $G_{r-2}/G_{r-1} \cong S_{r-1}$ and G_{r-1}, the possibilities for G_{r-2}. This is an example of **Hölder's Extension Problem**, which we can state in the following general form: Given two groups K and Q, determine all possible G such that $K \trianglelefteq G$ and $G/K \cong Q$. Such groups G are called **extensions** of K by the factor Q. In the language of group homomorphisms, the problem is to determine all groups G having a homomorphism f whose image is Q and whose kernel is K. (Notice that if a simple group G is an extension of K by Q, then either $G \cong K$ or $G \cong Q$. This property is the basis for a general definition of a simple algebraic system in terms of homomorphisms. See for example Curtis [**5**, p. 81]). That G need not be uniquely determined by K and Q is easily seen from the fact that both S_3 and $\boldsymbol{Z}_6$ (the cyclic group of order 6) are extensions of $\boldsymbol{Z}_3$ by $\boldsymbol{Z}_2$. The question is, can all possible such extensions G be constructed in some systematic way? The answer is yes, although not very economically. Schreier [**81**] in the early 1920's developed a technique for constructing all possible multiplication tables for G, but to date no general decision procedure is

available to identify which multiplication tables are those for isomorphic groups. (See Rotman [**34**, Chap. 7] or Scott [**36**, Chap. 9] for a more complete discussion of this.) Thus the list of possible multiplication tables will in general be plagued by repetitions, but still all possible extensions G of K by Q can be constructed.

Now that we see how to determine all the possible groups G_{r-2}, we can continue using Schreier's technique to construct all possible G_{r-3}, G_{r-4}, etc. At the top of the resulting tree will be all possible groups G which could correspond to the given list of composition factors. By considering all possible lists of composition factors, we would then survey all finite groups. Hence we see that the finite simple groups are the fundamental building blocks for the finite groups, and if we were able to classify all the finite simple groups, then we would be able to determine all possible finite groups (at least in the sense of obtaining their multiplication tables) by using the above scheme. The problem of classifying all finite simple groups is thus seen to be a fundamental and useful one in finite group theory.

2. The simple groups discovered up to 1955. A standard exercise in an undergraduate course in abstract algebra is to prove that the cyclic groups Z_p of prime order p are simple. These are the only abelian finite simple groups. Galois essentially showed that the alternating groups $A_n (n \geqq 5)$ constitute an infinite family of non-abelian finite simple groups. (For an elementary proof of simplicity see Jacobson [**28**, Vol. 1, p. 139].)

The next infinite families of finite simple groups were discovered among the **classical groups**, so named by Hermann Weyl in his book [**41**] published in 1939. These are the groups of matrices first introduced by Jordan [**32**] and whose structure was largely worked out by Dickson [**22, 51**] at the turn of this century. Dickson's methods were mostly *ad hoc* and highly computational, with much more elegant and readable treatments coming quite some time later in the work of Artin [**10, 43, 44**], Dieudonné [**21**], Huppert [**27**, Chap. II], and the new text of Jacobson [**31**, Chap. VI].

The first family of classical groups is the family of **general linear groups GL(V)** consisting of all automorphisms (i.e., invertible linear transformations) of a vector space V over a field K. The group $GL(V)$ is isomorphic (in many ways, one for each choice of a basis of V over K) to the group $GL_n(K)$ consisting of all n by n invertible matrices over K where n is the dimension of V over K. The **special linear groups** $\boldsymbol{SL(V)}$ (resp. $\boldsymbol{SL_n(K)}$) consist of the automorphisms of V (resp. matrices over K) which have determinant 1. Except in the case when $n = 2$ and K is the field of two elements, $SL_n(K)$ is the commutator subgroup of $GL_n(K)$. (Recall that the commutator subgroup G' of a group G is the subgroup generated by all commutators $xyx^{-1}y^{-1}$ as x and y vary through G, and that G/N is abelian if and only if $G' \subseteq N$.)

The remaining classical groups are the automorphism groups of certain non-degenerate forms on V. A **form** on V is a mapping $f: V \times V \to K$ such that $f(x_1 + x_2, y) = f(x_1, y) + f(x_2, y)$ and $f(x, y_1 + y_2) = f(x, y_1) + f(x, y_2)$ for all x, y, x_i, y_i in V. A form f is **non-degenerate** if its left and right kernels $L = \{x \in V \mid f(x, y) = 0 \text{ for all } y \text{ in } V\}$ and $R = \{y \in V \mid f(x, y) = 0 \text{ for all } x \text{ in } V\}$ are both zero (i.e., consist of only the zero vector of V). The **automorphism group** of a form f is the set of all automorphisms T of V which preserve the form in the sense that $f(Tx, Ty) = f(x, y)$ for all x, y in V. A form f is **symmetric bilinear** if $f(y, x) = f(x, y)$ and $f(\lambda x, y) = \lambda f(x, y)$ for all λ in K and all x, y in V. The automorphism group of such a form is called an **orthogonal group** and is denoted $\boldsymbol{O_n(K, f)}$. A form f is **skew-symmetric bilinear** if $f(y, x) = -f(x, y)$ and $f(\lambda x, y) = \lambda f(x, y)$ for all λ in K and all x, y in V. The automorphism group of such a form is called a **symplectic group** and is denoted $\boldsymbol{Sp_n(K, f)}$. If the characteristic of K is two then the definitions given above for orthogonal groups and symplectic groups coincide and it turns out that the groups one usually calls orthogonal groups must preserve a quadratic form in addition to the symmetric bilinear form. For this reason some of the statements below about orthogonal groups need refinement when the characteristic of K is two. (For further details consult Chevalley [**18**, Chap. 1].) The only other forms of interest here are Hermitian (sesquilinear) forms. Here K must be a separable quadratic extension of a field F such that there is an automorphism $a \to \bar{a}$ of K over F of order two. (The best-known example of this is K the complex

field, F the real field, and $\bar{a}$ the complex conjugate of a.) A form f is **Hermitian** (sesquilinear) if $f(y, x) = \overline{f(x, y)}$ and $f(\lambda x, y) = \lambda f(x, y)$ for all λ in K and all x, y in V. The automorphism group of such a form is called a **unitary group** and is denoted $U_n(K, f)$. The remaining classical groups (i.e., in addition to $GL_n(K)$ and $SL_n(K)$ described above) are then the automorphism groups of non-degenerate forms of these three types.

Remarks: Skew-symmetric forms are non-degenerate only when n is even [**10**, p. 119], say $n = 2m$. Further, any two non-degenerate forms on V give isomorphic symplectic groups, so the f in the notation is dropped and we write $Sp_{2m}(K)$. All the elements of $Sp_{2m}(K)$ have determinant 1 [**10**, p. 139], and $Sp_{2m}(K)$ coincides with its commutator subgroup, except when $m = 1$ and K has only two or three elements or $m = 2$ and K has only two elements [**10**, p. 173]. For $m = 1$ the groups $Sp_2(K)$ are isomorphic to the corresponding groups $SL_2(K)$ defined above. If the field K is finite (we remind the reader that we are studying finite groups), then any two non-degenerate symmetric bilinear forms on V give isomorphic orthogonal groups if n is odd (in which case f may again be omitted in the notation). But there are precisely two isomorphism classes of orthogonal groups if n is even according to whether the form has maximal or non-maximal Witt index. (We omit details here and refer the interested reader to Artin [**10**, Chap. 3].) The commutator subgroup of $O_n(K, f)$ is denoted $\Omega_n(K, f)$, and is in general a proper subgroup of the rotation group $SO_n(K, f)$ consisting of the elements of determinant 1 in $O_n(K, f)$. Again, if K is finite then any two non-degenerate Hermitian forms on V have isomorphic automorphism groups, so the notation is abbreviated to $U_n(K)$. For n greater than two, excepting only the case $n = 3$ and K a field of four elements, the commutator subgroup of $U_n(K)$ coincides with the subgroup $SU_n(K)$ of elements of determinant 1 in $U_n(K)$. For $n = 2$ the groups $SU_2(K)$ are isomorphic to the groups $SL_2(K)$ above.

There is a uniform procedure which yields simple groups from the classical groups. First form the commutator subgroup G' and then form the quotient group G'/C' of G' modulo its center C'. (This center consists of scalar multiples of the identity and so is not a complicated group. For instance, if K is finite, then C' is isomorphic to some subgroup of the cyclic multiplicative group of K.) The main structure theorem for the classical groups states that in most cases G'/C' is a simple group. There are a number of low dimensional exceptions over small fields and further exceptions occur in the unitary and orthogonal groups corresponding to **anisotropic** forms. (A form f is anisotropic if $f(x, x) \neq 0$ for all $x \neq 0$, a property we recognize as holding for the ordinary Euclidean inner product in $\boldsymbol{R}^n$.) For the orthogonal groups the structure theorem still holds for many important fields, and for a discussion in greater detail, see Dieudonné [**21**, Chap. II, §12]. This structure theorem, despite its general nature, was proved, even in the elegant treatments of Artin [**10**] and Dieudonné [**21**], by considering separately the various families of classical groups. In some cases (notably when K has characteristic two in the orthogonal groups) the work involved was quite formidable, and thus a general proof seemed all the more valuable. However, the intrinsic differences between the various cases posed a very stubborn barrier until the work of Chevalley (discussed in the next section) provided a uniform proof for all cases in addition to a much better understanding of the classical groups. No matter how it is proved, the main structure theorem does provide six families of groups whose members are simple in most cases. These are $PSL_n(K) = SL_n(K)/\text{Center}$, $PSp_{2m}(K) = Sp_{2m}(K)/\text{Center}$, $PSU_n(K) = SU_n(K)/\text{Center}$, and three families of orthogonal groups $\Omega_n(K, f)/\text{Center}$, two families arising when n is even, one family when n is odd. (For more details see Carter [**4**] or Jacobson [**31**, Chap. VI].) If K is a finite field, then these groups are finite simple groups, and, despite some isomorphisms between members of different families, the distinct groups comprise six doubly infinite families.

Dickson also found families of simple groups related to the simple Lie algebras of type G_2 and E_6 over the complex field (see Section 3 below) in three papers [**52, 53, 54**] which appeared shortly after his book [**22**].

Up to 1955 no further finite simple groups were known, except for five apparently sporadic (in contrast to the infinite families just discussed) groups discovered by Mathieu in 1861 and 1873 [**74, 75**].

These were formerly considered to be rather bizarre simple groups of no real importance in the general classification theory, but recent developments have increased their prominence considerably and in Section 5 below we shall describe them and indicate their importance in contemporary finite simple group theory.

We have now mentioned all the finite simple groups which were known up to 1955, and in fact essentially all were already known in 1905. Relatively little progress was made in classifying finite simple groups during this entire half century, an extraordinarily long period in contemporary mathematical history. In 1955, two papers appeared which radically changed this situation, one almost immediately, the other more gradually. The rest of this paper is largely a recounting of the after effects of these two papers.

3. The role of Lie algebras. Lest it be assumed that any attempt to classify all simple members of a category of algebraic structures is hopelessly ambitious, the example of Lie algebras affords a counterexample. In 1894 E. Cartan **[46]** completed the classification first begun by Killing **[71]** of the simple finite dimensional Lie algebras over C. The close analogy between finite dimensional Lie algebras and finite groups, at least in basic terms (for example, solvable, nilpotent, and simple are defined in essentially identical fashion in both theories), suggested that a classification of finite simple groups might well be attainable, and might even resemble in some ways the classification of simple Lie algebras given by Cartan. In this section we describe the latter classification briefly and then discuss the important application finite dimensional simple Lie algebras over C have found in the study of finite simple groups.

A **Lie algebra** L over the complex field is an algebra (a vector space with a bilinear product denoted $[x, y]$) for which the identities $[x, x] = 0$ and $[[x, y], z] + [[y, z], x] + [[z, x], y] = 0$ hold. An **ideal** I of L is a vector subspace satisfying $[x, I] \subseteq I$ for all $x \in L$. (This implies $[I, x] \subseteq I$ also, in view of the anti-commutativity of multiplication which follows directly from $[x, x] = 0$ and consideration of $[x + y, x + y]$.) L is **simple** if $[L, L] \neq 0$ and the only ideals of L are 0 and L. An example of a simple Lie algebra over the complex field is $sl(2, C)$, the set of 2 by 2 matrices of trace 0 under the commutator product $[A, B] = AB - BA$. This algebra has dimension 3 with

$$e = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad f = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \text{and} \quad h = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

comprising a basis. This example, while not deep, is illustrative of two important general facts. First, any associative algebra can be used to produce a Lie algebra if the commutator product is defined as above. Second, just as in the case of groups, a number of families of simple Lie algebras arise directly from matrices. Moreover, the classical groups over the complex field are **Lie groups** (see Cohn **[19]** or Chevalley **[17]**) and their simple subgroups described in Section 2 correspond in a canonical way to the simple Lie algebras of matrices just alluded to. Thus more than simple analogy links simple groups and simple Lie algebras.

The classification of Cartan rests on the systems of roots of the simple Lie algebras. For complete details of this the reader is referred to the books of Bourbaki **[12, 13]**, Carter **[16]**, Humphreys **[25]**, Jacobson **[29, 30]**, Samelson **[35]**, Serre **[39]**, and Winter **[42]**. First the linear map $\operatorname{ad} x$ is defined for any $x \in L$ by $\operatorname{ad} x(y) = [x, y]$. The **Killing form** K is a nondegenerate symmetric bilinear form on L defined by $K(x, y) = \operatorname{Trace}(\operatorname{ad} x \circ \operatorname{ad} y)$ for $x, y \in L$ where Trace is the familiar trace of a linear transformation. The **roots** of a Lie algebra L are certain linear mappings of H into C, where H is a **Cartan subalgebra**, a nilpotent subalgebra satisfying $[x, H] \subseteq H$ only when $x \in H$.

More precisely, a **root** of L relative to H is a linear map $r : H \to C$ for which there is an element $y_r \neq 0$ in L (called a **root vector** belonging to r) which satisfies $\operatorname{ad} h(y_r) = r(h)y_r$ for all $h \in H$. As an example, in $sl_2(C)$ the diagonal matrices $\operatorname{diag}(\alpha, -\alpha)$ of trace 0 comprise a Cartan subalgebra. If t_1 and t_2 are the linear maps which send $\operatorname{diag}(\alpha, -\alpha)$ to α and $-\alpha$ respectively, then $r = t_1 - t_2$ is a root of $sl_2(C)$ relative to this Cartan subalgebra. A corresponding root vector is $\left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$. Notice that a root

vector is a simultaneous eigenvector for all the linear transformations ad h as h ranges over H, with each eigenvalue given by the value of the root at h.

It can be shown that K remains nondegenerate when restricted to the Cartan subalgebra H, and this permits H to be identified with its dual space H^* under the correspondence $t \leftrightarrow h_t$, where h_t is the unique element [**28**, Vol. II, p. 141] of H satisfying $t(x) = K(h_t, x)$ for all $x \in H$. Thus one can transfer the Killing form to H^* by defining $K(t, s) = K(h_t, h_s)$. Over C, H^* is spanned by the roots. If we consider the rational subspace E_Q of H^* spanned by the set of roots, it turns out that K has only rational values on E_Q and is positive definite. If we then extend the scalars to $\boldsymbol{R}$ we get a Euclidean space E with inner product K. E has a basis consisting of **simple roots**, that is, roots r_i such that every root r is an integral linear combination of the r_i, with the integral coefficients either all positive integers or all negative integers. Using the Euclidean inner product in a clever way, one can reduce the description of the possible simple Lie algebras to a classification of the possible connected **Dynkin diagrams**. These diagrams have $\dim H$ vertices with the ith vertex joined to the jth by $n_{ij} = 4(r_i, r_j)^2/(r_i, r_i)(r_j, r_j)$ edges, with an arrow pointing toward the ith vertex if and only if $(r_i, r_i) < (r_j, r_j)$. That n_{ij} is always a non-negative integer (in fact is 0, 1, 2, or 3) is a fundamental tool in proving that the only possible connected Dynkin diagrams are those listed in Figure 1.

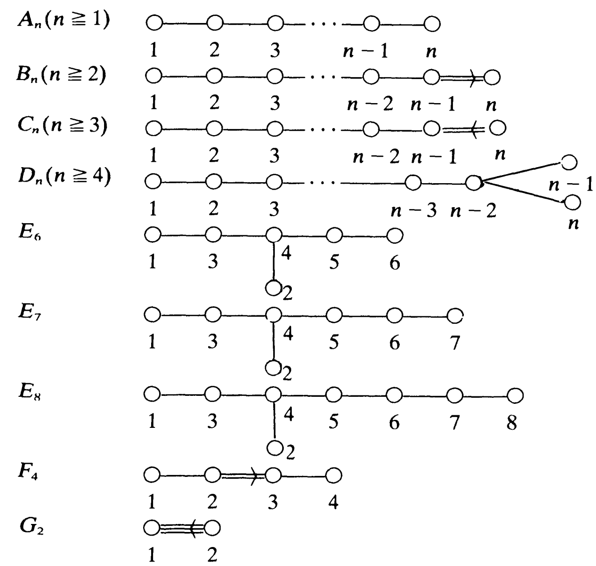


FIG. 1

The first Dynkin diagram A_n corresponds to the Lie algebra $sl_{n+1}(C)$ and the next three correspond to the classical simple Lie algebras of matrices which are related to the orthogonal and symplectic groups. These four infinite families of finite dimensional simple algebras thus play a role analogous to the infinite families of finite simple groups. The five individual exceptional algebras may call to mind the five Mathieu groups. Are the finite simple groups perfectly analogous to these Lie algebras in some mysterious way, and is this the reason no new finite simple groups were discovered after 1905? Had they all been found already?

No, they hadn't. In a paper now recognized as historic, Chevalley [**47**] developed a procedure for constructing infinite families of finite simple groups (called **Chevalley groups**) corresponding to each of the simple finite dimensional Lie algebras over C. Before giving a description of what Chevalley did, let us see how we might view the finite groups $SL_n(Z_p)$ (p a prime) as arising from $SL_n(C)$. First, $SL_n(C)$ can be thought of as imbedded in $M_n(C)$, the vector space of all n by n matrices with complex entries. In $M_n(C)$ consider a lattice Λ (i.e., a finitely generated additive subgroup whose C-span is

$M_n(C)$) which is also closed under matrix multiplication. While it is not obvious at first glance that such a Λ exists, a little reflection shows that $M_n(Z)$, the set of n by n matrices with integer entries, affords an example (and by no means the only one). Now $\Lambda \cap SL_n(C)$ gives us a multiplicative group. If we use $M_n(Z)$ for Λ then the group is $SL_n(Z)$. When reduced modulo p, this latter group yields $SL_n(Z_p)$, which is customarily written $SL_n(p)$.

Chevalley in essence generalized this construction by showing how to carry it out in any finite dimensional simple Lie algebra L over C instead of in $M_n(C)$. More precisely, he constructed a lattice $\Lambda = L_Z$ generated by a basis B of L consisting of certain elements h_i in the Cartan subalgebra H, and certain root vectors e_r corresponding to the non-zero roots r of L relative to H. With respect to B, the constants of structure (those scalars which arise when a product of two basis elements is written as a linear combination of basis elements) are all integers, so L_Z is closed under multiplication. Similarly, if K is any field, the free K-module L_K on B is closed under multiplication and so is a Lie algebra over K since the (integer) constants of structure can always be interpreted as elements of K. Now a classical procedure involving the exponential series produces from L a group H analogous to $SL_n(C)$ above. Chevalley showed how to carry out this procedure using L_K in place of L. This refines and extends the idea above of reducing $\Lambda \cap SL_n(C)$ modulo p, and produces a group $G(L, K)$ which is simple (except when $L = sl_2(C)$ and K has only two or three elements or L is of type B_2 or G_2 and K has only two elements). In particular, if K is finite then $G(L, K)$ is a finite simple group. For instance, if $L = sl_n(C)$ and $K = Z_p$, then $G(L, K) = PSL_n(p)$.

The Chevalley group $G(L, K)$ bears essentially the same relation to L_K that the complex Lie group $G(L, C)$ bears to its Lie algebra L. So in addition to its importance to group theory, Chevalley's paper also opened up new directions by providing a purely algebraic analogue for Lie groups. (See for example Humphreys [**26**], Borel [**11**], Seligman [**37**], or Steinberg [**9**, **40**].)

Ree [**78**] quickly showed that Chevalley's procedure produced the classical simple groups described in Section 2, except for the unitary groups and some of the orthogonal groups and also produced the groups found by Dickson which correspond to L of type G_2 and E_6. Chevalley himself had shown that the infinite families corresponding to L of type F_4, E_7 and E_8 were new. Chevalley's methods, moreover, provided a single argument that proved simplicity simultaneously in all cases, thereby avoiding the kind of case-by-case analysis formerly employed in the study of the classical groups.

Variations on Chevalley's procedure soon provided the remaining classical groups and also further new infinite families of finite simple groups. The first such construction was discovered by R. Steinberg [**83**] (and independently by J. Tits [**86**, **87**] and D. Hertzig [**61**]), and is based on the fact that certain Dynkin diagrams (namely, A_n, $n \geqq 2$, D_n, $n \geqq 4$, and E_6) possess symmetries which lead in a natural way to automorphisms of L. For instance, in type D_n the transposition which interchanges nodes $n-1$ and n gives rise to an automorphism of L of order 2. Steinberg showed that if K is a field which has an automorphism of the same order as the diagram automorphism of L, then these two automorphisms compose to yield an automorphism of L_K which when suitably combined with Chevalley's construction produces a subgroup $G^1(L, K)$ of $G(L, K)$ called a **twisted Chevalley group**. $G^1(L, K)$ is simple in every case except when K has four elements and $L = sl_3(C)$. The remaining families of classical simple groups are obtained in this way, as well as two new infinite families of simple groups. The proof of simplicity closely resembles Chevalley's, and Tits [**88**] has shown how to abstract the argument to produce a uniform proof of simplicity for the classical simple groups and the additional simple groups of Lie type.

In 1960 Suzuki [**84**], working outside Lie theory, constructed a new infinite family of simple groups. Ree showed that these groups correspond to a more complicated graph automorphism of the Lie algebras of type B_2, and proceeded to construct similar groups corresponding to algebras of type G_2 and F_4 ([**79**] and [**80**]; a good exposition is given in Carter [**4**] or [**16**, pp. 223ff.]).

The techniques of Lie theory were thus applied with most impressive results to the problem of classifying the finite simple groups. In the early 1960's many group theorists felt that a full

classification of the finite simple groups was at hand, and would consist of a general characterization of the simple groups constructible from Lie algebras.

4. Classification problems. For the reasons outlined at the end of the Introduction (as well as for other reasons) we would like to have a description or classification of the finite simple groups. For instance, the use of an external device, Lie theory, which gives a uniform description of all the classical matrix groups as well as some others (including Dickson's groups of type G_2 and E_6) is a step in this direction. But it has shortcomings, on the one hand in that it does not encompass all the simple groups, and on the other hand in that it does not give a purely group theoretic description of the simple groups. In this section we describe some of the directions such group theoretic approaches have taken.

In its most general form the problem might be posed as follows: Given a set S with an associative (partial) multiplication defined on it, determine all simple groups G having a subset isomorphic to S with this (partial) multiplication. This problem is far too general, so one might impose some additional conditions on S. To begin with, one might insist that S be a group and thus a subgroup of any finite simple group G containing it. But even this is much too general. For example, if we choose S to be a group isomorphic to A_5, the alternating group of degree five, then the possibilities for G include all the alternating groups $A_n (n \geqq 5)$, infinitely many of the groups $PSL_2(q)$, and many others. In this case we have given all the information one could possibly want about S itself, since we have specified S as a group up to isomorphism. Therefore, the only hope of cutting down the list of possibilities for G is to provide some information about how S is embedded in G. This can be done in many ways, of which a few are: Requiring that S be a maximal subgroup of G, requiring that S be a Sylow p-subgroup of G for some prime p, requiring that S be the centralizer or normalizer in G of some subset of G, or requiring that S be the stabilizer of a point in some special type of permutation representation of G. The basic principle is as follows: Since S is a proper subgroup of G (excepting one case where S itself is a simple group), its composition factors have order smaller than that of G and thus are (at least in principle) less complicated groups than G, and the information about how S is embedded in G is used as a filter to pick out a reasonable (perhaps even finite) number of simple groups G containing S as a subgroup in this particular manner. The first significant result in this direction came in the second momentous paper of 1955 mentioned at the end of Section 2:

THEOREM (Brauer and Fowler [**45**]). *Suppose G is a simple group and t is an involution (i.e., an element of order two) of G. If $S = C_G(t) = \{x \in G \mid xt = tx\}$ (i.e., the* ***centralizer*** *of t in G), then* $|G| < |S|^2!$.

The importance of this theorem is that it tells us that if a group S is the centralizer of an involution in some simple group G, then there are only finitely many possibilities for G since the order of G is bounded as a function of that of S. In addition, it suggests the **centralizer of an involution problem**:

Given a group S containing an involution in its center, determine all finite simple groups G such that $S \cong C_G(t)$ for some involution t in G.

In a first course in abstract algebra one often encounters the elementary exercise of proving that a group of even order contains an involution. In 1963, Feit and Thompson [**55**] proved that every non-cyclic finite simple group must have even order, and so must contain an involution. This rein forces the naturalness of attempting to classify finite simple groups by centralizers of involutions as proposed by Richard Brauer at the 1954 International Congress of Mathematicians. Brauer not only posed this problem, but also initiated the work on it. It was quickly found that some choices of S gave rise to no possibilities for G, others to a unique group G, and yet others to several groups G. In the beginning one frequently chose S to be isomorphic to the centralizer of an involution in some known group G_0 in the hopes of characterizing G_0 as the only possible group G having S as the centralizer of an involution. However, almost at the outset there were surprises. Brauer, in attempting to characterize $G_0 = PSL_3(3)$ ($= SL_3(3)$), took S to be $GL_2(3)$, and found, in addition to the expected

answer $G = SL_3(3)$, the unexpected solution $G = M_{11}$, where M_{11} is one of the Mathieu groups to be discussed in the next section. As experience was gained with the centralizer of an involution problem, an important theme emerged. This was that one wanted to use the centralizer of an involution to determine the Sylow 2-subgroups of G and then use this to determine the group or groups G. This suggested the following problem:

Given a 2-group S determine all simple groups G having Sylow 2-subgroups isomorphic to S.

Unlike the centralizer of an involution problem, this problem need not have a finite set of solutions. For instance, if $S \cong \mathbf{Z}_2 \times \mathbf{Z}_2$ (the direct product of two cyclic groups of order two), then there are infinitely many possibilities for G, namely the groups $PSL_2(q)$ with $q \geqq 5$ and $q \equiv \pm 3 \bmod(8)$, (Dickson **[22]**, Chapter XII). Nevertheless, it seems that specifying a Sylow 2-subgroup gives a very direct route to the groups G, and even in the cases where there are infinitely many such groups G they seem to appear in a definite pattern which is fairly easy to describe and understand. Of course, many of the choices for a Sylow 2-subgroup S give rise to no possible groups G or to a unique group G.

Thus far characterization problems in which S is a subgroup have been discussed. Another type of characterization, the study of which was initiated by Fischer, is to assume that S is a conjugacy class of elements (usually involutions) of G, which generate G, and where assumptions are made about the subgroups of G generated by small subsets of S, in particular two-element subsets. Fischer's original problem was to determine the groups G generated by a conjugacy class D of involutions having the property that the product of any two members of this class has order at most three. In this setting, if the product has order one (i.e., is the identity) then the two involutions are actually one and the same involution and the group generated by them is isomorphic to $\mathbf{Z}_2$. If the product has order two, then they are commuting involutions and the group they generate is isomorphic to $\mathbf{Z}_2 \times \mathbf{Z}_2$. Finally, if the product has order three then they generate a symmetric group of degree three. Fischer was able to solve this problem under the additional assumptions that $\mathrm{Sol}(G)$, the largest normal solvable subgroup of G, is trivial, and that the second commutator group G'' of G coincides with the commutator subgroup G' of G. The solution set consists of finite groups G such that G' is simple and includes, besides several infinite families, three new simple groups which will be described in some detail in Section 6. Since then, Fischer and some of his students have gone on to investigate the same problem under the assumption that products of order at most four occur for two involutions in the class. Although this work is not yet complete it has already led to the discovery of three more new simple groups as well as the discovery of another possible simple group for which existence is not yet settled.

In working on a characterization problem one accumulates a wealth of information about any finite simple group G satisfying the hypotheses of the problem. In many cases this information is inconsistent, so there is no such group. Otherwise it is possible to determine the order of such a group (or in some cases several possible orders for several solutions to the problem). If this order is the order of some known finite simple group G_0, then one must still prove that G must be isomorphic to G_0 and this is usually not difficult. (However, there is one case, namely the characterization by the centralizer of an involution of Ree's simple groups of type G_2, where Ward **[90]** showed that any group G satisfying the hypotheses has the same order and shares many other properties with the corresponding Ree group, but as yet it has not been proved that G must be a Ree group. This is in spite of the attention this particular problem has received due to its being the only gap in the determination of the finite simple groups with abelian Sylow 2-subgroups.) If, on the other hand, this order is not the order of any known simple group, there remains the problem of proving the existence (and uniqueness) of a finite simple group of that order satisfying the hypotheses of the characterization problem. The tool which is most frequently used at this stage is representation theory, a branch of algebra which has a long and illustrious history of successes in solving a wide variety of problems including a central role in Feit and Thompson's proof of the solvability of groups of odd order.

Because representation theory itself is not germane to this paper we say only as much about it as seems necessary to explain how it is used in proving the existence of finite simple groups. (For more information consult Brauer [3], Curtis & Reiner [20], or Feit [23].) A **representation** of a group H is a homomorphism ρ mapping H into the group $GL(V)$ of automorphisms of a vector space V over a field K. If V is n-dimensional over K, then $GL(V)$ is isomorphic to the group $GL_n(K)$ of n by n invertible matrices over K, so ρ may also be viewed as a homomorphism of H into $GL_n(K)$ and we use these points of view interchangeably. The integer n is called the **degree** of ρ. A representation ρ of H is **irreducible** if no proper subspace of V is left invariant by the image $\rho(H)$ of H in $GL(V)$. Two representations ρ_1 and ρ_2 are **equivalent** if $\rho_2 = \sigma\rho_1\sigma^{-1}$ for some element σ of $GL(V)$. This is an extension of the familiar concept of similarity of matrices. A representation over a field of characteristic zero (resp. not zero) is called **ordinary** (resp. **modular**). If H is a finite group, then H has only a finite number of inequivalent irreducible ordinary representations and any ordinary representation of H is a direct sum of irreducible representations. Associated with each ordinary representation, ρ is the **character** of ρ, a function $\chi_\rho : H \to K$ defined by $\chi_\rho(x) = \mathrm{Tr}(\rho(x))$ for each x in G where Tr is the familiar trace of a matrix. Characters are important because two ordinary representations of H are equivalent if and only if they have the same character. If K is algebraically closed of characteristic zero (e.g., C), then the characters of the inequivalent irreducible representations of H over K define an array called the **character table** of H. There exist many arithmetic relationships among the entries of a character table and the character table of a group provides a wealth of information about its group theoretic properties including information about its subgroups. The character table of a group also provides information about the modular representations of the group via the modular characters, which are complex valued functions that play for modular representations a role similar to that of characters for ordinary representations. This is important because many groups have modular representations (over fields of characteristic dividing the group order) with degrees much smaller than those of any non-trivial ordinary representation. We are now ready to establish the connection with the rest of this section.

The wealth of information about a group G satisfying the hypotheses of some characterization problem is used not only to find the order of the group but also its character table. It is a tribute to the advanced state of the art in representation theory that it is possible to construct the character table of a group from as little as a knowledge of its order, the assumption that it is simple, and perhaps some purely group theoretic information about the group and *without even knowing if the group exists.* In some cases this is the point at which a contradiction to the existence of any such group G is reached, as the accumulated information about G may be inconsistent with any possible character table of G. On the other hand, having a character table does not yet guarantee the existence of G. But it is a good beginning in that no characterization problem has ever gotten as far as the construction of a character table without it subsequently being established that the predicted group G exists. As the above example of the Ree groups shows, uniqueness is a more delicate problem. Assuming one has a character table of G, the most obvious thing to do is to choose a character corresponding to a representation of G of relatively small degree n and use this to construct G as a group of n by n matrices. Essentially this idea was used to construct the Rudvalis group and the Thompson group, the values of n in the respective cases being 28 and 248. For many groups the calculations involved in trying to construct them as groups of matrices are too involved, so other ideas must be used. The character table can also be used to find information about the subgroups of G which can then be used to construct G as a group of permutations on the set of cosets of such a subgroup. Essentially this idea was used to construct two of Janko's groups, the Held group, the Lyons group, the O'Nan group, and Fischer's "baby monster" group. The idea of using the character table to find the modular characters, and then using them to construct G as a group of matrices over a finite field was used by Janko [66] to construct the first of his groups as a group of 7 by 7 matrices over Z_{11}. (If he had used ordinary representations, he would have had to use at least 56 by 56 matrices.)

In this section we have described some of the main themes of the group theoretic approach to

classifying the finite simple groups, but have, apart from a few passing references, left out some of the most interesting products of this approach, the sporadic simple groups, so that they could be discussed in a roughly historical order later. In Section 5 we take up the Mathieu groups which were discovered in the mid-late nineteenth century and in Section 6 we describe the sporadic simple groups which have been discovered more recently, among them the groups mentioned in the preceding paragraph and Fischer's groups referred to earlier.

5. The Mathieu Groups. Before the Mathieu groups can be discussed properly, some terminology from the theory of permutation groups must be introduced. A **permutation representation** of a finite group G is a homomorphism $\rho: G \to S(X)$ of G into the symmetric group $S(X)$ consisting of all the permutations of some set X. The representation ρ is called **faithful** if the kernel of ρ consists of only the identity element of G. The image of a permutation representation, or more generally any subgroup of $S(X)$, is called a **permutation group** on the set X. If H is a permutation group on X, then for each $x \in X$ the **stabilizer** H_x of x is the set $\{h \in H \mid hx = x\}$ and it is easy to see that H_x is a subgroup of H. Given a permutation group H on X, an equivalence relation $\sim$ on X is defined by $x \sim y$ means $y = hx$ for some element h of H. The equivalence classes of $\sim$ are called the **orbits** of H on X or simply H-orbits. If X itself is an H-orbit, then H is said to be a **transitive** permutation group. If $x, y \in X$ are in the same H-orbit, then the stabilizers H_x and H_y are conjugate subgroups of H, so that in particular all stabilizers are conjugate in H when H is a transitive permutation group on X.

A permutation group G on a set X is said to be ***k*-ply transitive** if any ordered k-tuple of distinct elements of X can be mapped onto any other k-tuple of distinct elements of X by a permutation in G. If $k = 1$, then the group G is merely *transitive* on X, for all that is asserted is that any element of X can be mapped to any other by some permutation in G. If $k \geqq 2$, then G is said to be *multiply transitive*. The symmetric group S_n of degree n, consisting of all the permutations of a set with cardinality n, is k-ply transitive for any integer k $(k \leqq n)$, while the alternating group A_n, consisting of all the even permutations of the same set, is k-ply transitive for any integer k $(k \leqq n-2)$. The symmetric groups S_n (for $n \geqq 6$) and the alternating groups A_m (for $m \geqq 8$) are the only known permutation groups which are k-ply transitive for any integer $k \geqq 6$. One of the outstanding problems in the theory of permutation groups is to determine whether these are the only examples. In the other direction, there are infinitely many examples of doubly and even triply transitive permutation groups which are not symmetric or alternating groups. The automorphism groups of projective geometries over finite fields are always doubly transitive on points and in the case of projective lines they are even triply transitive on points. (See Carmichael [**15**, Chapter 12].)

Other than the symmetric and alternating groups, there are only four known permutation groups which are 4-ply or 5-ply transitive on any finite set. These are the **Mathieu groups** which were first described by Mathieu in two papers of 1861 [**74**] and 1873 [**75**]. The Mathieu group M_{12} is a group of order $95040 = 2^6 3^3 5 \cdot 11$ which acts 5-ply transitively on a set of cardinality 12, and the Mathieu group M_{24} is a group of order $244{,}823{,}040 = 2^{10}3^3 5 \cdot 7 \cdot 11 \cdot 23$ which acts 5-ply transitively on a set of cardinality 24. The stabilizers of any point in these respective groups are groups M_{11} of order $7920 = 2^4 3^2 5 \cdot 11$ and M_{23} of order $10{,}200{,}960 = 2^7 3^2 5 \cdot 7 \cdot 11 \cdot 23$ which act 4-ply transitively on sets of cardinality 11 and 23, respectively. The stabilizer of a point in M_{23} is the fifth Mathieu group M_{22} of order $443{,}520 = 2^7 3^2 5 \cdot 7 \cdot 11$ which acts 3-ply transitively on a set with cardinality 22. (The stabilizer of a point in M_{11} is a group of automorphisms of the projective line over the field with nine elements, while the stabilizer of a point in M_{22} is a group of automorphisms of the projective plane over the field with four elements. Both of these groups, or more precisely their simple composition factors, are groups appearing in the Lie type classification of Section 3.) The five groups M_{11}, M_{12}, M_{22}, M_{23}, and M_{24} are all simple and are not isomorphic to any of the alternating groups or to any of the groups of Lie type. Since they are not known to belong to any infinite families of simple groups, Burnside [**14**, Note N, p. 504] referred to them as *sporadic simple groups* and this term is now used for any simple group which does not belong to any of the infinite families described in Sections 2 and 3.

In order to give a reasonable description of the Mathieu groups, we follow Witt **[91]** and introduce the idea of a (k, l, m)-**Steiner system**. A (k, l, m)-Steiner system is a collection of l-element subsets (referred to as ***l*-clubs**) of a set X with cardinality m having the property that each k-element subset of X belongs to a unique l-club. (We are interested only in nontrivial Steiner systems, i.e., ones with $k < l$.) Projective planes and higher dimensional projective spaces over finite fields provide infinitely many examples of $(2, l, m)$-Steiner systems if one takes X to be the set of all projective points, and lets the set of l-clubs be the set of all projective lines. The desired property is then assured by the fact that any two projective points determine a unique projective line. There are also many $(3, l, m)$-Steiner systems, but the only known $(4, l, m)$-Steiner systems have parameters $(4, 5, 11)$ and $(4, 7, 23)$, and the only known $(5, l, m)$-Steiner systems have parameters $(5, 6, 12)$ and $(5, 8, 24)$. There are no known (k, l, m)-Steiner systems with $k \geqq 6$. The above four Steiner systems are unique up to isomorphism in that any (k, l, m)-Steiner system with the same parameters is essentially one of the four above.

The **automorphism group** A of a (k, l, m)-Steiner system is the set of all those permutations of the set X which preserve the set of l-clubs in the sense that every l-club is mapped to an l-club by every permutation in A. For instance, in the geometric examples above A is the collineation group of the projective geometry. The automorphism groups of the $(5, 6, 12)$- and $(5, 8, 24)$-Steiner systems are the previously mentioned groups M_{12} and M_{24}, respectively. The automorphism groups of the $(4, 5, 11)$- and $(4, 7, 23)$-Steiner systems are the groups M_{11} and M_{23}, respectively. The group M_{22} is a subgroup of index two in the automorphism group of the $(3, 6, 22)$-Steiner system derived from the $(4, 7, 23)$-Steiner system by restriction. This $(3, 6, 22)$-Steiner system is also unique up to isomorphism. The process of restriction works as follows: Any (k, l, m)-Steiner system gives rise to a $(k-1, l-1, m-1)$-Steiner system by deletion of any element from X and taking as the new $(l-1)$-clubs all those l-clubs which contain the deleted element.

The number of l-clubs in a (k, l, m)-Steiner system is easily seen to be $\binom{m}{k}/\binom{l}{k}$ because X has $\binom{m}{k}$ k-element subsets, each l-club has $\binom{l}{k}$ k-element subsets, and each k-element subset of X belongs to a unique l-club. For instance, the $(5, 8, 24)$-Steiner system has

$$\binom{24}{5}\Big/\binom{8}{5} = 24\cdot 23\cdot 11\cdot 7/8\cdot 7 = 759$$

8-clubs (or as they are more often called, **octads**), while the $(5, 6, 12)$-Steiner system has $\binom{12}{5}/\binom{6}{5} = 132$ 6-clubs (or **hexads**). It is relatively easy to show that each of the Mathieu groups acts transitively on the set of l-clubs of its corresponding Steiner system, but we will not go into this matter here.

Rather surprisingly, the Mathieu groups also arise in information theory as the automorphism groups of certain perfect error correcting codes over $GF(3)$ and $GF(2)$. These so called Golay codes are the only perfect codes which correct more than one error. The interested reader is referred to two excellent survey articles **[1, 2]** by Assmus and Mattson for further details. The action of the Mathieu groups on these codes is also important for the study of simple groups since this action plays a role in the construction of several of the sporadic simple groups to be mentioned in the following section.

6. Sporadic simple groups. The remaining known (as of Fall 1976) sporadic simple groups, with the exception of the three Conway groups, were discovered either in the course of work on a classification problem (as described in Section 4) or as permutation groups of a very special type, namely rank 3 permutation groups. The three Conway groups were discovered as automorphism groups of a lattice in 24-dimensional space that was first described by Leech and which appears to owe its existence to the existence of the $(5, 8, 24)$-Steiner system.

In the course of classifying simple groups with abelian Sylow 2-subgroups, one must consider the following classification problem: Determine the simple groups G having (a) Sylow 2-subgroups isomorphic to $Z_2 \times Z_2 \times Z_2$, and (b) an involution t such that $C_G(t) \cong Z_2 \times PSL_2(q)$. Janko and Thompson **[70]** showed that if $q > 5$, then $q = 3^n$ for some integer n. Then a theorem of Ward **[90]** showed that G has the same order as the Ree group ${}^2G_2(3^n)$ mentioned in Sections 3 and 4. Janko **[66]**

then went on to consider the case $q = 5$ (in which case $PSL_2(5) \cong A_5$) and was able to prove that any such group J must have order $2^3 3 \cdot 5 \cdot 7 \cdot 11 \cdot 19$ and a unique character table. Janko then investigated the modular characters of J and was able to prove that J has a modular irreducible representation of degree 7 over $GF(11)$. He was then able to prove existence and uniqueness of J by exhibiting two 7×7 matrices over $GF(11)$ which generate the group. (For a while it was thought that J might be the first member of a family of groups of order $q(q^3-1)(q+1)$ with $q = 11^n$ for suitable n [**4**, p. 238], but so far no such family has ever been found and it seems unlikely that there is any such family.)

Janko's discovery was a mixed blessing for the researchers working on the classification problems as it shattered the hope that the only simple groups were the cyclic and alternating groups, the groups of Lie type, and the five sporadic Mathieu groups. On the other hand, it provided a new simple group as a direct consequence of research on a particular classification problem. The possibility that more new groups might be found spurred renewed research activity in finite simple groups, and before long many more sporadic simple groups were found.

The next two sporadic simple groups were found by an examination of **rank 3 permutation groups**. A transitive permutation group G on a set X is said to be **rank *m*** if the stabilizer G_x of any point x has exactly m orbits on X (or equivalently, if G has exactly m orbits on ordered pairs of elements of X). So in a rank 3 group, G_x has $\{x\}$ as an orbit and two other oribts which are denoted $\Delta(x)$ and $\Gamma(x)$ where the notation is such that $\Delta(gx) = g\Delta(x)$. A rank 3 group thus has the strongest transitivity property short of being multiply transitive. D. Higman [**62**] investigated rank 3 groups combinatorially and introduced the notation $|\Delta(x)| = k$, $|\Gamma(x)| = l$, and $|\Delta(x) \cap \Delta(y)| = \lambda$ or μ according as y is in $\Delta(x)$ or $\Gamma(x)$. He was able to show that $\mu l = k(k - \lambda - 1)$, that certain more complicated relations between the parameters k, l, λ, and μ hold, and that the degrees of the two nontrivial irreducible constituents of the permutation character of (G, X) could be calculated from these parameters. The conditions on the parameters are such that with the aid of an electronic computer a parameter list with all possible parameter sets such that $|X| = k + l + 1 < 200$ could easily be written down.

Higman found particularly fascinating the parameter set $(22, 77, 0, 6)$ because it suggested the possibility of a permutation group G on a set of cardinality $100 = 1 + 22 + 77$ with $G_x \cong M_{22}$ or $\mathrm{Aut}(M_{22})$ and where the G_x-orbits of order 22 and 77 are the points and hexads (6-clubs) of the $(3, 6, 22)$-Steiner system. D. Higman and Sims [**63**] constructed a graph Ω whose 100 vertices were the 22 points ($= |\Delta(x)|$) and 77 hexads ($= |\Gamma(x)|$) of a $(3, 6, 22)$-Steiner system together with a new vertex x. The vertex x was joined to each of the 22 points in $\Delta(x)$ and to none of the 77 hexads in $\Gamma(x)$. Each point y in $\Delta(x)$ was joined to x, to none of the other points in $\Delta(x)$ (because $\lambda = 0$), and to the 21 hexads containing the point y. Each hexad z in $\Gamma(x)$ was not joined to x, was joined to the 6 points contained in z, and to the 16 hexads which are disjoint from z. Higman and Sims then used Witt's proof [**91**] of the uniqueness of the $(3, 6, 22)$-Steiner system to prove that the graph had a transitive automorphism group, which must be rank 3 by the nature of the construction of Ω. The automorphism group of Ω is not simple but has a simple commutator subgroup of index two and order $2^9 3^2 5^3 7 \cdot 11$ which is not isomorphic to any of the previously known simple groups. This group is now known as the *Higman–Sims group* (abbreviation Hi–S).

The next sporadic simple group was discovered by Hall and Wales [**57**] as a rank 3 group although its existence had been predicted by Janko [**67**]. Janko was studying certain classification problems and was able to show that if G is a simple group in which the centralizer of an involution was a nonabelian group of order 2^5 acted upon faithfully by A_5, then G has either one or two conjugacy classes of involutions. In the first case $|G| = 2^7 3^5 5 \cdot 17 \cdot 19$ and in the second case $|G| = 2^7 3^3 5^2 7$. It was the second of these possibilities which Hall and Wales proved to exist. They first showed that if such a group G exists, then G must have a rank 3 permutation representation of degree 100 in which $G_x \cong PSU_3(3)$ and the parameters are $(36, 63, 14, 12)$. Thereafter they actually constructed permutations on a set of cardinality 100 which generated a simple group of the right order and they were also able to prove the uniqueness of this group which is now known as the *Hall–Janko–Wales group* (abbreviated Ha–J–W).

Thereafter new rank 3 sporadic simple groups were found by McLaughlin [**76**] and Suzuki [**85**]. Since McLaughlin knew that the group $PSU_4(3)$ had subgroups of index 112 and 162, the parameter set (112, 162, 30, 56) suggested to him the possible existence of a rank 3 permutation group G of degree 275 with $G_x \cong PSU_4(3)$. Using ideas similar to those of Higman and Sims he constructed a graph with 275 vertices and showed it to have a transitive automorphism group. The automorphism group of this graph has a simple commutator subgroup of index two and order $2^7 3^6 5^3 7 \cdot 11 = 275 \,|\, PSU_4(3) \,|$. This group is now known as the *McLaughlin group* (abbreviation McL). Suzuki also studied rank 3 groups and he asked if there might be a rank 3 permutation group H in which $H_x \cong$ Ha–J–W and in which one of the H_x orbits had order 100. It turned out that there was a unique possibility with parameters (100, 315, 36, 20), the H_x-orbit of order 315 corresponding to the centralizer of an involution studied by Janko. Suzuki constructed the appropriate graph and showed that it had a transitive automorphism group. Unfortunately this group, or more precisely, its commutator subgroup, turned out to be the known Dickson group $G_2(4)$, rather than a new sporadic simple group. Undeterred, Suzuki then asked if there might be a rank 3 group G such that $G_x \cong G_2(4)$ and where one of the G_x-orbits was the above permutation group of degree $416 = 1 + 100 + 315$. Again it turned out that there was a unique possibility with parameters (416, 1365, 100, 96), and again Suzuki constructed the appropriate graph and showed that it had a transitive automorphism group. The automorphism group of this graph had a simple commutator subgroup of index two and order $2^{13} 3^7 5^2 7 \cdot 11 \cdot 13 = 1782 \cdot 416 \,|$ Ha–J–W $|$. This group is now known as the *Suzuki group* (abbreviated Sz).

The group of order $2^7 3^5 5 \cdot 17 \cdot 19$ predicted by Janko was more difficult to construct, as its character table (also determined by Janko) showed that it did not have any low rank permutation representations. Existence was eventually proved by Graham Higman and McKay [**64**] who gave a presentation of the group by generators and relations, and then used a computer to verify that this defined the group predicted by Janko. This group is now known as the *Higman–Janko–McKay group* (abbreviated H–J–McK).

The next three simple groups were discovered by Conway [**48**] in his study of the automorphism group of the lattice in 24-dimensional space which had been constructed by Leech [**72**] in connection with sphere packing and covering problems in higher dimensional spaces. The automorphism group A of this lattice is not a simple group but has a center Z consisting of the 24×24 identity matrix and its negative. The quotient group A/Z is a simple group of order $2^{21} 3^9 5^4 7^2 11 \cdot 13 \cdot 23$ and this group is called ·1. Among the composition factors of the stabilizers of sublattices Conway found two more new simple groups, ·2 of order $2^{18} 3^6 5^3 7 \cdot 11 \cdot 23$, and ·3 or order $2^{10} 3^7 5^3 7 \cdot 11 \cdot 23$. Curiously, many other sporadic simple groups appear in various ways in A. All the Mathieu groups appear since they, or more precisely, their associated Steiner systems, were instrumental in Leech's construction of the lattice. The groups Hi–S and McL appear among the composition factors of the stabilizers of sublattices, while the groups Ha–J–W and Sz appear among the composition factors of the centralizers of certain elements of order 5 and 3 respectively in A. The remaining groups J and H–J–McK cannot appear because their orders do not divide that of A.

Held [**59**] systematically studied the problem of classifying the Mathieu groups M_{22}, M_{23}, and M_{24} by the centralizers of involutions. Although he found only the expected answers in the cases of M_{22} and M_{23}, in the case of M_{24} he was able to show [**60**] that if G is a simple group with the centralizer of an involution isomorphic to the centralizer of an involution in the center of some Sylow 2-subgroup of M_{24}, then one of the following must occur: 1) G is isomorphic either to M_{24} or to $PSL_5(2)$ (these are the expected answers), 2) G has order $2^{10} 3^3 5^2 7^3 17$. In the latter case Held was able to determine all the conjugacy classes of elements of G and Thompson was then able to construct the character table of G. Actually proving the existence of the group was again accomplished by Graham Higman and McKay [**65**] who determined a presentation for the group and verified using a computer that the presentation defined a group of the right order and having the predicted properties. This group is now known as the *Held group* or the *Held–Higman–McKay group* and the abbreviations He and H–H–M are used.

The next three sporadic simple groups were discovered by Fischer [56] during his investigation of the following classification problem: Determine the groups G having a conjugacy class D of involutions such that G is generated by D and such that the product of any two elements of D has order at most three. This problem was suggested by the conjugacy class of transpositions in the symmetric groups so Fischer called this the problem of groups generated by **3-transpositions**. This problem turned out to be clearly related to the work on rank 3 permutation groups because the permutation representation of G acting by conjugation on the elements of D is rank 3 in all the interesting cases. Under certain technical assumptions (which force the commutator subgroups of G to be simple) Fischer was able to classify the groups generated by 3-transpositions. In addition to the expected families of groups: S_n (the symmetric groups for $n>4$), $Sp_{2n}(2)$, $O_{2n}(2,\pm)$, $O_n(3,\pm)$, and $PSU_n(2)$, three isolated groups M(22), M(23), and M(24) appeared. The first two of these were new simple groups of orders $2^{17}3^95^27\cdot 11\cdot 13$ and $2^{18}3^{13}5^27\cdot 11\cdot 13\cdot 17\cdot 23$ while the last one had a simple commutator subgroup M(24)′ of index two and order $2^{21}3^{16}5^27^311\cdot 13\cdot 17\cdot 23\cdot 29$ (or approximately 1.26×10^{24}). These groups are now known as the *Fischer groups.* The notation M(22), M(23), and M(24)′ was introduced by Fischer because of certain very strong connections between these groups and the corresponding Mathieu groups. In some sources the abbreviations Fi_{22}, Fi_{23}, and Fi_{24}' are used for these groups.

The next simple group was predicted by Lyons [73] as the result of work on a classification problem. The centralizer H of any involution t in McL is such that the quotient group $H/\langle t\rangle$ is isomorphic to A_8 (the alternating group of degree eight). Thompson posed the problem of classifying the simple groups G containing an involution t such that $C_G(t)/\langle t\rangle\cong A_n$ and t is in the commutator subgroup of $C_G(t)$ and succeeded in proving that if $n\geqq 12$ then there is no such group G. The cases $n\leqq 8$ had been settled earlier; Janko [68] proved that there was no such group for $n=9$; so the remaining cases $n=10$ and 11 were given by Thompson to his student Lyons as research problems. Lyons showed that there was no such group for $n=10$, and for $n=11$ he showed that any such group must have order $2^83^75^67\cdot 11\cdot 31\cdot 37\cdot 67$ and constructed its character table. The information from the character table was not very promising as far as settling existence of the group was concerned, for it showed that the group had no faithful matrix representations of degree less than 2480 in any field of characteristic 0, nor any permutation representation of degree less than eight million. It seemed at the outset that proving existence of this group would be a very difficult problem, but Sims [82] eventually proved existence by using a computer to construct generating permutations on a set of cardinality about eight million and verifying that these permutations generated a group of the right order satisfying all of Lyons' conditions. This group is now known as the *Lyons–Sims group* and the abbreviations LyS or simply Ly are used.

Thereafter a period of several years passed during which no new simple groups were found or predicted. Rudvalis, one of the authors here, during a systematic investigation of rank 3 groups in which the action on one of the orbits for the stabilizer of a point is isomorphic to the action of that group on a conjugacy class of involutions, predicted the existence of a rank 3 group G in which the stabilizer of a point is the Ree group ${}^2F_4(2)$—this group is not simple but has a simple commutator subgroup of index two—mentioned in Section 3. With some help from Feit and Lyons, Rudvalis and Frame were able to determine the character table of G and from this it followed that G had a projective representation of degree 28 over the field $Q(i)$ of Gaussian numbers. Using this information Conway and Wales [50] proved existence of the group by exhibiting matrices which generated this projective representation. This group has order $2^{14}3^35^37\cdot 13\cdot 29$ and is known as the *Rudvalis group.* The abbreviations R and R–C–W are used.

The existence of yet another simple group was predicted by O'Nan [77] as the result of work on a classification problem. Existence again proved difficult to establish but eventually, as in the case of LyS, Sims was able to use the computer to construct permutations which generated the group. This group has order $2^93^45\cdot 7^311\cdot 19\cdot 31$ and is known as the *O'Nan group.* The abbreviation O'N is used, as is O'N–S.

TABLE I. The Simple Groups of Lie Type

Name or Discoverer	Lie Notation	Order of G	d
$PSL_n(q)$ $(n \geqq 2)$	$A_{n-1}(q)$	$(1/d)q^{n(n-1)/2}\prod_{i=2}^{n}(q^i-1)$	$(n, q-1)$
$PSU_n(q)$ $(n \geqq 3)$	${}^2A_{n-1}(q)$	$(1/d)q^{n(n-1)/2}\prod_{i=2}^{n}(q^i-(-1)^i)$	$(n, q+1)$
$PSp_{2n}(q)$ $(n \geqq 2)$	$C_n(q)$	$(1/d)q^{n^2}\prod_{i=1}^{n}(q^{2i}-1)$	$(2, q-1)$
$P\Omega_{2n+1}(q)$ $(n \geqq 3)$	$B_n(q)$	$(1/d)q^{n^2}\prod_{i=1}^{n}(q^{2i}-1)$	$(2, q-1)$
$P\Omega_{2n}(q,+)$ $(n \geqq 4)$	$D_n(q)$	$(1/d)q^{n(n-1)}(q^n-1)\prod_{i=1}^{n-1}(q^{2i}-1)$	$(4, q^n-1)$
$P\Omega_{2n}(q,-)$ $(n \geqq 4)$	${}^2D_n(q)$	$(1/d)q^{n(n-1)}(q^n+1)\prod_{i=1}^{n-1}(q^{2i}-1)$	$(4, q^n+1)$
Steinberg	${}^3D_4(q)$	$q^{12}(q^8+q^4+1)(q^6-1)(q^2-1)$	
Dickson	$G_2(q)$	$q^6(q^6-1)(q^2-1)$	
Chevalley	$F_4(q)$	$q^{24}(q^{12}-1)(q^8-1)(q^6-1)(q^2-1)$	
Dickson	$E_6(q)$	$(1/d)q^{36}(q^{12}-1)(q^9-1)(q^8-1)(q^6-1)(q^5-1)(q^2-1)$	$(3, q-1)$
Steinberg	${}^2E_6(q)$	$(1/d)q^{36}(q^{12}-1)(q^9+1)(q^8-1)(q^6-1)(q^5+1)(q^2-1)$	$(3, q+1)$
Chevalley	$E_7(q)$	$(1/d)q^{63}(q^{18}-1)(q^{14}-1)(q^{12}-1)(q^{10}-1)(q^8-1)(q^6-1)(q^2-1)$	$(2, q-1)$
Chevalley	$E_8(q)$	$q^{120}(q^{30}-1)(q^{24}-1)(q^{20}-1)(q^{18}-1)(q^{14}-1)(q^{12}-1)(q^8-1)(q^2-1)$	
Suzuki $Sz(q)$ $(q=2^{2n+1})$	${}^2B_2(q)$	$q^2(q^2+1)(q-1)$	
Ree $R_1(q)$ $(q=3^{2n+1})$	${}^2G_2(q)$	$q^3(q^3+1)(q-1)$	
Ree $R_2(q)$ $(q=2^{2n+1})$	${}^2F_4(q)$	$q^{12}(q^6+1)(q^4-1)(q^3+1)(q-1)$	

Comments on Table I. In addition to the families of simple groups of Lie type there are the cyclic groups Z_p, of prime order p, and the alternating groups A_n $(n \geqq 5)$, of order $(n!/2)$. In the above table q is an arbitrary power of an arbitrary prime, except in the last three cases where (as indicated) it must be an odd power of two or three. Also (n, s) in the definition of d stands for the greatest common divisor of n and s. The groups G in the table are non-abelian simple groups except in the following cases:

$PSL_2(2)$, $PSL_2(3)$, $PSU_3(2)$, and ${}^2B_2(2)$ are all solvable groups.
$PSp_4(2)$, $G_2(2)$, and ${}^2F_4(2)$ all have a simple commutator subgroup G' of index two in G.
${}^2G_2(3)$ has a simple commutator subgroup G' of index three in G.

Some families have been omitted from the table due to the isomorphisms:

$$\mathbf{B}_1(\mathbf{q}) \cong \mathbf{C}_1(\mathbf{q}) \cong \mathbf{A}_1(\mathbf{q}) \cong {}^2\mathbf{A}_1(\mathbf{q}) \qquad \mathbf{B}_2(\mathbf{q}) \cong \mathbf{C}_2(\mathbf{q})$$

$$\mathbf{D}_2(\mathbf{q}) \cong \mathbf{A}_1(\mathbf{q}) \times \mathbf{A}_1(\mathbf{q}) \qquad {}^2\mathbf{D}_2(\mathbf{q}) \cong \mathbf{A}_1(\mathbf{q}^2)$$

$$\mathbf{D}_3(\mathbf{q}) \cong \mathbf{A}_3(\mathbf{q}) \qquad {}^2\mathbf{D}_3(\mathbf{q}) \cong {}^2\mathbf{A}_3(\mathbf{q}).$$

For fields of characteristic two (i.e., $q = 2^m$): $B_n(q) \cong C_n(q)$ $(n \geqq 3)$.
For individual groups there are the following isomorphisms:

$$A_5 \cong PSL_2(4) \cong PSL_2(5) \qquad A_6 \cong PSL_2(9) \cong PSp_4(2)' \qquad A_8 \cong PSL_4(2)$$

$$PSL_2(7) \cong PSL_3(2) \qquad PSp_4(3) \cong PSU_4(2) \qquad PSU_3(3) \cong G_2(2)' \qquad PSL_2(8) \cong {}^2G_2(3)'.$$

We observe in passing that the groups $PSp_{2n}(q)$ have the same order as the groups $P\Omega_{2n+1}(q)$ but the corresponding groups are not isomorphic for $n \geqq 3$ and $q \neq 2^m$, so there are infinitely many pairs of non-isomorphic simple groups of the same order. The only other known instance of this is that A_8 and $PSL_3(4)$ both have order $2^6 3^2 5 \cdot 7 = 20,160$.

TABLE II. Sporadic Simple Groups

Name	Date of Discovery	Order
M_{11}	1861 by Mathieu	$2^4 3^2 5 \cdot 11 = 7920$
M_{12}	1861 by Mathieu	$2^6 3^3 5 \cdot 11 = 95040$
M_{22}	1873 by Mathieu	$2^7 3^2 5 \cdot 7 \cdot 11 = 443{,}520$
M_{23}	1873 by Mathieu	$2^7 3^2 5 \cdot 7 \cdot 11 \cdot 23 = 10{,}200{,}960$
M_{24}	1873 by Mathieu	$2^{10} 3^3 5 \cdot 7 \cdot 11 \cdot 23 = 244{,}823{,}040$
J or Ja	1965 by Janko	$2^3 3 \cdot 5 \cdot 7 \cdot 11 \cdot 19 = 175{,}560$
Hi-S	1967 by D. Higman & Sims	$2^9 3^2 5^3 7 \cdot 11 = 44{,}352{,}000$
Ha-J-W or J_2	1967 by Hall & Wales	$2^7 3^3 5^2 7 = 604{,}800$
McL	1968 by McLaughlin	$2^7 3^6 5^3 7 \cdot 11 = 898{,}128{,}000$
Sz or Suz	1968 by Suzuki	$2^{13} 3^7 5^2 7 \cdot 11 \cdot 13 = 448{,}345{,}497{,}600$
H-J-McK or J_3	1968 by G. Higman & McKay	$2^7 3^5 5 \cdot 17 \cdot 19 = 50{,}232{,}960$
$\cdot 1$ or Co_1	1968 by Conway	$2^{21} 3^9 5^4 7^2 11 \cdot 13 \cdot 23 = 4{,}157{,}771{,}806{,}543{,}360{,}000$
$\cdot 2$ or Co_2	1968 by Conway	$2^{18} 3^6 5^3 7 \cdot 11 \cdot 23 = 42{,}305{,}421{,}312{,}000$
$\cdot 3$ or Co_3	1968 by Conway	$2^{10} 3^7 5^3 7 \cdot 11 \cdot 23 = 495{,}766{,}656{,}000$
He or H-H-McK	1968 by Held, G Higman & McKay	$2^{10} 3^3 5^2 7^3 \cdot 17 = 4{,}030{,}387{,}200$
M(22) or Fi_{22}	1969 by Fischer	$2^{17} 3^9 5^2 7 \cdot 11 \cdot 13 = 64{,}561{,}751{,}654{,}400$
M(23) or Fi_{23}	1969 by Fischer	$2^{18} 3^{13} 5^2 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 23 = 4{,}089{,}470{,}473{,}293{,}004{,}800$
M(24)′ or Fi_{24}'	1969 by Fischer	$2^{21} 3^{16} 5^2 7^3 11 \cdot 13 \cdot 17 \cdot 23 \cdot 29 = 1{,}255{,}205{,}709{,}190{,}661{,}721{,}292{,}800$
Ly or Ly-S	1970 by Lyons-Sims	$2^8 3^7 5^6 7 \cdot 11 \cdot 31 \cdot 37 \cdot 67 = 51{,}765{,}179{,}004{,}000{,}000$
R or R-C-W	1972 by Rudvalis-Conway-Wales	$2^{14} 3^3 5^3 7 \cdot 13 \cdot 29 = 145{,}926{,}144{,}000$
O'N or O'N-S	1973 by O'Nan-Sims	$2^9 3^4 5 \cdot 7^3 11 \cdot 19 \cdot 31 = 460{,}815{,}505{,}920$
T	1974 by Thompson-Smith	$2^{15} 3^{10} 5^3 7^2 13 \cdot 19 \cdot 31 = 90{,}745{,}943{,}887{,}872{,}000$
Ha-C-N-S	1974 by Harada-Conway-Norton-Smith	$2^{14} 3^6 5^6 7 \cdot 11 \cdot 19 = 273{,}030{,}912{,}000{,}000$
F or F-L-S	1973 by Fischer (existence 1976 by Leon-Sims)	$2^{41} 3^{13} 5^6 7^2 11 \cdot 13 \cdot 17 \cdot 19 \cdot 23 \cdot 31 \cdot 47 \approx 4.15 \times 10^{34}$
		Possible Groups
M	1974 proposed by Fischer	$2^{46} 3^{20} 5^9 7^6 11^2 13^3 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71 \approx 8.08 \times 10^{53}$
J_4	1975 proposed by Janko	$2^{21} 3^3 5 \cdot 7 \cdot 11^3 23 \cdot 29 \cdot 31 \cdot 37 \cdot 43 = 86{,}775{,}571{,}046{,}077{,}562{,}880$

Additional simple groups have come from Fischer's preliminary work on the classification of the groups generated by a conjugacy class D of involutions with the property that the product of any two elements of D has order at most four. This work suggested the existence of simple groups of orders

$$2^{41} 3^{13} 5^6 7^2 11 \cdot 13 \cdot 17 \cdot 19 \cdot 23 \cdot 31 \cdot 47 \text{ and } 2^{46} 3^{20} 5^9 7^6 11^2 13^3 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71.$$

These two groups are referred to as "Fischer's monsters." The first is called "baby monster" and sometimes abbreviated F, while the second is called simply "monster" and sometimes abbreviated M. The order of the latter group is about 10^{54} which does indeed make it monstrous. Although Leon and Sims have very recently (Fall 1976) announced the construction of "baby monster" as a permutation group, the existence of "monster" has not been settled as of 1976. Assuming M exists, then it has elements x_3 and x_5 of orders 3 and 5, respectively, such that the centralizers $C_M(x_i)$ are isomorphic to the direct products $\langle x_i \rangle \times G_i$, where G_3 is a simple group of order $2^{15} 3^{10} 5^3 7^2 13 \cdot 19 \cdot 31$ and G_5 is a simple group of order $2^{14} 3^6 5^6 7 \cdot 11 \cdot 19$. The existence of these two groups has been proved (independently of the existence of M) by *Thompson* **[89]** in the first case and by the work of *Harada* **[58]** and *Conway*, *Norton* and *Smith* **[49]** in the second. Thompson (with help from P. Smith) exhibited his group as a group of automorphisms of a lattice in 248-dimensional space which was derived from

$E_8(\boldsymbol{C})$ and also as a subgroup of the Chevalley group $E_8(\boldsymbol{Z}_3)$. Harada investigated the other group by the centralizer of an involution method, and among other things, determined its character table. This information was used by Conway, his research student S. Norton, and P. Smith to investigate certain subgroups of the group and eventually Norton and Smith found generating permutations for the group using a computer.

In June 1975 Janko predicted the existence of a new simple group as the solution of a centralizer of an involution characterization problem [**69**]. The character table of this group has been worked out, but as yet (December 1976) existence has not been proved. The group has order $2^{21}3^35\cdot7\cdot11^323\cdot29\cdot31\cdot37\cdot43$ and presumably it will be referred to as the *fourth Janko group* (abbreviation J_4) until such a time as existence is proved at which point it may be renamed after Janko and the person(s) who prove its existence.

7. Concluding Remarks. The discovery of the post-Mathieu sporadic simple groups described in Section 6 has certainly made it appear that the optimistic point of view expressed at the end of Section 3 was at best premature and at worst perhaps fundamentally erroneous. Few mathematicians are left today who believe that we stand on the threshold of a complete classification of the finite simple groups. But the fact that the earlier visualization of the scope of the finite simple groups seems to have been too narrow does not in itself mean that a complete classification is beyond the realm of possibility or likelihood. In this section we discuss some conjectures about the question we started with: Can the class of finite simple groups be satisfactorily catalogued?

A first conjecture that might be made is that all the infinite families of finite simple groups have been found, and that the class of sporadic groups is finite. In this view, the sporadic groups comprise a finite number of exceptional groups, perhaps 25, perhaps 50, perhaps 500, but a finite number all of which will be discovered in a matter of years or decades. One cannot yet write this possibility off, essentially the view which was widely held up until the last five to ten years. But its appeal has diminished considerably in the wake of the discovery of the new sporadic simple groups. Basic mathematical and human instincts seem to discourage belief in exceptional behavior which is too widespread. With each new sporadic group another exception must be accepted by a holder of this view, and many former holders have abandoned it.

A second possibility is that there are still one or more additional infinite families of which the sporadic groups provide a few examples. A number of the sporadic groups described above are seen to be rather closely related to groups of Lie type, which suggests that some of the infinite families of groups of Lie type might not yet be complete. Perhaps more sophisticated Lie techniques could produce further infinite families. At present there is little evidence to support this conjecture.

Either of the possibilities just mentioned anticipates that the class of finite simple groups will in essence resemble the class of simple finite dimensional Lie algebras over the complex field. While the analogy between finite groups and finite dimensional Lie algebras is striking in a number of respects, it is far less clear that the analogy between simple groups and simple Lie algebras *over the complex field* is decisive. The class of simple finite dimensional Lie algebras over fields of characteristic not zero is still very much undetermined (see Seligman [**38**]). In view of this, it may be overly optimistic to expect a classification of finite simple groups that is as uncomplicated as that given in Section 3 for finite dimensional Lie algebras over C.

In any case, mathematicians believe that the objects they study can be described in an orderly manner, and the work of describing the class of finite simple groups goes forward on the two fronts of discovery of new examples (cf. Section 6) and classification of the simple groups in terms of certain conditions of a general group-theoretic nature (cf. Section 4). A great many positive classification theorems have been obtained, and in Gorenstein [**24**, Part III, **7**, and **8**] can be found a more detailed introduction than we have been able to give here to the methods growing out of Brauer's pioneering work. In Feit [**6**, Section 4] is a report on the status of this program up to 1970 with explicit lists of

classification theorems and references to a very complete bibliography. More classification results are continually appearing, and this constitutes a very active aspect of the subject.

Everything we have said so far in this section is based on the optimistic expectation that an orderly global description of the class of finite simple groups is possible. A less comfortable possibility is suggested by the fact that several of the new sporadic groups discussed in Section 6 as well as the Mathieu groups, discussed in Section 5, appear to result, at least in part, from certain numerical relationships of a possibly random nature, such as the existence of certain Steiner systems, or the existence of orbits of certain sizes. This possibility for the nature of the simple groups can be suggested perhaps by the nature of the first objects to which we compared the finite simple groups, the prime positive integers. The distribution of the primes among the positive integers is not describable in an ordered way, except locally in intervals between two integers. If the finite simple groups include infinitely many essentially randomly distributed sporadic groups, then the classification procedures just alluded to might go on without ever coming to a conclusion.

It is certainly a remarkable circumstance in which possibilities as disparate as those discussed in this section can be put forth about the ultimate resolution of a leading mathematical problem, especially a problem on which so much progress seems to have been made in the last two decades. The uniform Lie theoretic techniques which have produced the classical simple groups and new infinite families of finite simple groups, together with so much group theoretic information about them, and the body of classification results proved by group-theoretic techniques do stand as two impressive achievements of recent mathematics. It now appears that the continuing search for an answer to which (if any) of the possibilities discussed just above is most nearly accurate will remain an actively pursued problem for some time. But a mathematician in 1950 who could then have foreseen the outline of developments over the next twenty-five years could easily have called it quite formidable and something that would take a much longer time than it did. In any event, the answer to the question we raised at the start of this section, which may well depend on discoveries still to be made and theory still to be developed, seems likely to be a goal whose vigorous pursuit should maintain the study of finite simple groups as one of the most interesting and exciting research areas in modern algebra.

1st author's work supported in part by NSF Grant OIP75–08849.
2nd author's work supported in part by NSF Grant MPS73–0863–A02.

References

The references are divided into the categories listed to assist those interested in reading further about finite simple groups. Other expository articles are a natural place to begin, with textbooks then consulted for the topics needed to read the research articles. *Mathematical Reviews* citations are included not only because they provide short synopses of the content of each article, but also because they can often help one decide if the subject of the article reviewed is appropriate to one's interests and background. A convenient compilation of reviews of survey and research articles is: D. Gorenstein, ed., *Reviews of Papers on Finite Groups*, American Mathematical Society, Providence, R. I., 1974.

I. Survey articles

1. E. Assmus, Jr. and H. F. Mattson, Jr., On tactical configurations and error correcting codes, J. Combinatorial Theory, 2 (1967) 243–257. MR 36 # 64.

2. ——— and ———, Coding and combinatorics, SIAM Review, 16 (1974) 349–388. MR 50 # 12432.

3. R. Brauer, Representations of finite groups (in Lectures on Modern Mathematics, Vol. I edited by T. Saaty), Wiley, New York, 1963, 133–175. MR 31 # 231.

4. R. Carter, Simple groups and simple Lie algebras, J. London Math. Soc., 40 (1965) 193–240. MR 30 # 4855.

5. C. Curtis, The classical groups as a source of algebraic problems, this MONTHLY (50th Anniversary Issue), 74 (1967) Part II, 80–91. MR 34 # 5940.

6. W. Feit, The current situation in the theory of finite simple groups, Proc. International Congr. of Mathematicians (Nice, 1970), Tome I, Gauthier-Villars, Paris, 1971, 55–93.

7. D. Gorenstein, Centralizers of involutions in finite simple groups, Finite Simple Groups (Proc. Instructional Conf. Oxford, 1969), Academic Press, London–New York, 1971, 66–134. (See also in the same volume related articles by J. Conway, E. Dade, and G. Higman.) MR 49 # 402.

8. ———, Finite simple groups and their classification, Israel J. Math., 19 (1974) 5–66. MR 50 # 13244.

9. R. Steinberg, Algebraic groups and finite groups, Illinois J. Math., 13 (1969) 81–86. MR 38 # 2207.

II. Books

10. E. Artin, Geometric Algebra, Interscience, New York, 1957. MR 18, p. 553.

11. A. Borel, Linear Algebraic Groups, Benjamin, New York, 1969. MR 40 # 4273.

12. N. Bourbaki, Groupes et Algèbres de Lie, Chap. I, 2nd ed., Hermann, Paris, 1971. MR 42 # 6159.

13. ———, Groupes et Algèbres de Lie, Chap. IV, V, et VI, Hermann, Paris, 1968. MR 39 # 1590.

14. W. Burnside, Theory of Groups of Finite Order, 2nd ed., Dover, New York, 1955 (originally published, Cambridge Univ. Press, New York, 1911). MR 16, p. 1086.

15. R. Carmichael, Introduction to the Theory of Groups of Finite Order, Dover, New York, 1956 (originally published, Ginn, New York, 1937). MR 17 p. 823.

16. R. Carter, Simple Groups of Lie Type, Wiley, New York, 1972.

17. C. Chevalley, Theory of Lie Groups, Vol. I, Princeton Univ. Press, Princeton, N. J., 1946. MR 7, p. 412.

18. ———, Algebraic Theory of Spinors, Columbia University Press, New York, 1954. MR 15, p. 678.

19. P. Cohn, Lie Groups, Cambridge Math. Tract No. 46, Cambridge Univ. Press, Cambridge, 1957. MR 21 # 2702.

20. C. Curtis and I. Reiner, Representation Theory of Finite Groups and Associative Algebras, Wiley-Interscience, New York, 1962. MR 26 # 2519.

21. J. Dieudonné, La Géométrie des Groupes Classiques, Troisème Éd., Ergebnisse der Math. N. F. Band 5, Springer-Verlag, Berlin–Heidelberg–New York, 1971. MR 17, p. 236, MR 28 # 1239, MR 46 # 9186.

22. L. Dickson, Linear Groups with an Exposition of Galois Field Theory, Dover, New York, 1958 (originally published, Teubner, Leipzig, 1901). MR 21 # 3488.

23. W. Feit, Characters of Finite Groups, Benjamin, New York, 1967. MR 36 # 2715.

24. D. Gorenstein, Finite Groups, Harper & Row, New York, 1968. MR 38 # 229.

25. J. Humphreys, Introduction to Lie Algebras and Representation Theory, Springer-Verlag, New York–Heidelberg–Berlin, 1972. MR 48 # 2197.

26. ———, Linear Algebraic Groups, Springer-Verlag, New York–Heidelberg–Berlin, 1975.

27. B. Huppert, Endliche Gruppen I, Springer-Verlag, Berlin–Heidelberg–New York, 1967. MR 37 # 302.

28. N. Jacobson, Lectures in Abstract Algebra, Three Volumes, Van Nostrand, Princeton–New York, 1951, 1953, 1964. MR 12, p. 794; MR 14, p. 837; MR 30 # 3087.

29. ———, Lie Algebras, Wiley-Interscience, New York, 1962. MR 26 # 1345.

30. ———, Exceptional Lie Algebras, Marcel Dekker, New York, 1971. MR 44 # 1707.

31. ———, Basic Algebra I, Freeman, San Francisco, 1974. MR 50 # 9457.

32. C. Jordan, Traité des Substitutions et des Équations Algébriques, Librairie Scientifique et Technique A. Blanchard, Paris, 1957 (originally published Gauthier–Villars, Paris, 1870). MR 19, p. 937.

33. I. MacDonald, The Theory of Groups, Oxford Univ. Press, Oxford, 1968. MR 39 # 5670.

34. J. Rotman, The Theory of Groups: An Introduction, Allyn & Bacon, Boston, 1968. MR 34 # 4338.

35. H. Samelson, Notes on Lie algebras, Van Nostrand Reinhold Math. Studies No. 23, Van Nostrand Reinhold, New York, 1969. MR 40 # 7322.

36. W. Scott, Group Theory, Prentice-Hall, Englewood Cliffs, N.J., 1964. MR 29 # 4785.

37. G. Seligman, Algebraic Groups, Yale Univ. Math. Dept., New Haven, 1965.

38. ———, Modular Lie Algebras, Ergebnisse der Math. und ihrer Grenzgebiete, Band 40, Springer-Verlag, New York–Berlin–Heidelberg, 1967. MR 39 # 6933.

39. J.-P. Serre, Algèbres de Lie Semi-Simples Complexes, Benjamin, New York, 1966. MR 35 # 6721.

40. R. Steinberg, Lectures on Chevalley Groups, Yale Univ. Math. Dept., New Haven, 1968.

41. H. Weyl, The Classical Groups, 2nd ed., Princeton Univ. Press, Princeton, NJ, 1946. MR 1, p. 42.

42. D. Winter, Abstract Lie Algebras, MIT Press, Cambridge, Mass., 1972. MR 48 # 11230.

III. Research articles

43. E. Artin, The orders of the linear groups, Comm. Pure Appl. Math., 8 (1955) 355–366. MR 17, p. 12.

44. ———, The orders of the classical simple groups, Comm. Pure Appl. Math., 8 (1955) 455–472. MR 17, p. 457.

45. R. Brauer and K. Fowler, Groups of even order, Ann. of Math., 62 (1955) 565–583. MR 17, p. 580.

46. E. Cartan, Sur la structure des groupes de transformations finis et continus, Thèse, Paris, 1894; 2nd Ed. Viubert, Paris, 1933; Oeuvres Complètes, Partie 1, Tome 1, Gauthier-Villars, Paris, 1952. MR 14, p. 343.

47. C. Chevalley, Sur certains groupes simples, Tôhoku Math. J., 7 (1955) 14–66. MR 17, pp. 457–8.

48. J. Conway, A group of order 8,315,533,613,086,720,000, Bull. London Math. Soc., 1 (1969) 79–88. MR 40 # 1470.

49. J. Conway, S. Norton and P. Smith, To appear.

50. J. Conway and D. Wales, The construction of the Rudvalis simple group of order 145,926,144,000, J. Algebra, 27 (1973) 538–548. MR 49 # 400.

51. L. Dickson, The theory of linear groups in an arbitrary field, Trans. American Math. Soc., 2 (1901) 363–394.

52. ———, A class of groups in an arbitrary realm connected with the configuration of the 27 lines on a cubic surface I, Quarterly J. Pure & Applied Math., 33 (1901) 145–173.

53. ———, A new system of simple groups, Math. Annalen, 60 (1905) 137–150.

54. ———, A class of groups in an arbitrary realm connected with the configuration of the 27 lines on a cubic surface II, Quarterly J. Pure & Applied Math., 39 (1908) 205–209.

55. W. Feit and J. Thompson, Solvability of groups of odd order, Pacific J. Math., 13 (1963) 755–1029. MR 29 # 3538.

56. B. Fischer, Finite groups generated by 3-transpositions I, II, III, Invent. Math., 131 (1971) 232–246 and to appear (Lecture Notes, Math. Institute, Univ. of Warwick, Coventry, 1970. MR 45 # 3557).

57. M. Hall, Jr., and D. Wales, The simple group of order 604,800, J. Algebra, 9 (1968) 417–450. MR 39 # 1544.

58. K. Harada, On the simple group F of order $2^{14}3^{6}5^{6}7\cdot 11\cdot 19$, Proc. of the Conference on Finite Groups (ed. W. Scott & F. Gross), Academic Press, New York, 1976, 119–276.

59. D. Held, Eine Kennzeichnung der Mathieu-Gruppe M_{22} und der alternierden Gruppe A_{10}, J. Algebra, 8 (1968) 436–449. MR # 5283.

60. ———, The simple groups related to M_{24}, J. Algebra, 13 (1969) 253–296. MR 40 # 2745.

61. D. Hertzig, Forms of algebraic groups, Proc. Amer. Math. Soc., 12 (1961) 657–660. MR 23 # A2486.

62. D. Higman, Finite permutation groups of rank 3, Math. Z., 86 (1964) 145–156. MR 32 # 4182.

63. D. Higman and C. Sims, A simple group of order 44,352,000, Math. Z., 105 (1968) 110–113. MR 37 # 2854.

64. G. Higman and J. McKay, On Janko's simple group of order 50,232,960, Bull. London Math. Soc., 1 (1969) 89–94. MR 40 # 224.

65. ——— and ———, A presentation for Held's group, unpublished.

66. Z. Janko, A new finite simple group with abelian Sylow 2-subgroups and its characterization, J. Algebra, 3 (1966) 147–186. MR 33 # 1359.

67. ———, Some new simple groups of finite order I, Symposia Mathematica (INDAM) Rome, 1967–8, Vol. I, Academic Press, London–New York, 1969, 25–64. MR 39 # 5686.

68. ———, The nonexistence of a certain type of finite simple group, J. Algebra, 18 (1971) 245–253. MR 44 # 2822.

69. ———, A new finite simple group of the order 86,775,571,046,077,562,880 which possesses M_{24} and the full cover of M_{22} as subgroups, J. Algebra, 42 (1976) 564–596.

70. Z. Janko and J. Thompson, On a class of finite simple groups of Ree, J. Algebra, 4 (1966) 274–292. MR 34 # 1386.

71. W. Killing, Die Zusammensetzung der stetigen endlichen Transformationgruppen I, II, III, IV, Math. Annalen, 31 (1888) 252–290; 33 (1889) 1–48; 34 (1889) 57–122; 36 (1890) 161–189.

72. J. Leech, Notes on sphere packings, Canad. J. Math., 19 (1967) 251–267. MR 35 # 878.

73. R. Lyons, Evidence for a new finite simple group, J. Algebra, 20 (1972) 540–569. MR 45 # 8722.

74. E. Mathieu, Mémoire sur l'étude des fonctions de plusieurs quantités, J. Math. Pures Appl., Series 2, 6 (1861) 241–243.

75. ———, Sur les fonctions cinq fois transitives des 24 quantités, J. Math. Pures Appl., Series 2, 18 (1873) 25–46.

76. J. McLaughlin, A simple group of order 898,128,000, Theory of Finite Groups (Symposium, Harvard Univ., 1968) ed. R. Brauer and C.-H. Sah, Benjamin, New York, 1969, 109–111, MR 39 # 4268.

77. M. O'Nan, Some evidence for the existence of a new simple group, to appear.

78. R. Ree, On some simple groups defined by C. Chevalley, Trans. Amer. Math. Soc., 84 (1957) 392–400. MR 19, p. 37.

79. ———, A family of simple groups associated with the simple Lie algebra of type F_4, Amer. J. Math., 83 (1961) 401–420. MR 23A # 2460b, MR 24A # 2617.

80. ———, A family of simple groups associated with the simple Lie algebra of type G_2, Amer. J. Math., 83 (1961) 432–462. MR 23A # 2460a.

81. O. Schreier, Über die Erweiterung von Gruppen I, Monatsh. Math. Phys., 34 (1926) 165–180.

82. C. Sims, The existence and uniqueness of Lyons' group, Finite Groups '72 (proc. Gainesville Conf., Univ. of Fla., Gainesville, Fla. 1972), North-Holland, Amsterdam, 1973, 138–141. MR 50 # 7358.

83. R. Steinberg, Variations on a theme of Chevalley, Pacific J. Math., 9 (1959) 875–891. MR 22 # 79.

84. M. Suzuki, On a class of doubly transitive groups I, II, Ann. of Math., 75 (1962) 105–145; 79 (1964) 514–589. MR 22 # 11038, MR 25 # 112.

85. ———, A simple group of order 448,345,497,600, Theory of Finite Groups (Symposium, Harvard Univ., 1968), Benjamin, New York, 1969, 113–119. MR 39 # 2867.

86. J. Tits, Les "formes réelles" des groupes de type E_6, Séminaire Bourbaki 10e année 1957/58, Exposé 162, Secrétariat Mathématique, Paris, 1958. MR 21 # 4981.

87. ———, Sur la trialité et certains groupes qui s'en déduisent, Inst. Hautes Études Scientifiques Publ. Math., 2 (1959) 13–60.

88. ———, Algebraic and abstract simple groups, Ann. of Math., 80 (1964) 313–329. MR 20 # 2259.

89. J. Thompson, A simple subgroup of $E_8(3)$, Abstracts of International Symposium of Theory of Finite Groups, Department of Mathematics, Hokkaido University, Sapporo, Japan, 1974.

90. H. Ward, On Ree's series of simple groups, Trans. Amer. Math. Soc., 121 (1966) 62–89. MR 33 # 5722.

91. E. Witt, Über Steinersche Systeme, Abh. Math. Sem. Univ. Hamburg, 12 (1938) 265–275.

Department of Mathematics, University of Connecticut, Storrs, CT 06268.

Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01002.

ADDENDUM TO "CAN ONE SEE THE SHAPE OF A SURFACE?"

(This Monthly, 84 (1977) 259-270)

Reese T. Prosser

I learned too late to include in the original article that recent work by Michael Taylor [3] and Andrew Majda [1, 2] has finally placed on a firm basis the heuristic arguments I gave for the inversion of convex boundary scattering at high frequencies. Using modern techniques of Fourier Integral Operators and delicate estimations in the time domain, they are able to show that my equation (43) is rigorously correct as stated.

References

1. Andrew Majda, Comm. Pure Appl. Math, 29 (1976) 261-291.

2. ———, Comm. Pure Appl. Math, 30 (1977).

3. Michael Taylor, Comm. Pure Appl. Math, 29 (1976) 1-38.

Department of Mathematics, Dartmouth College, Hanover, NH 03755.

PROGRESS REPORTS

Edited by P. R. Halmos

Material for this Department should be sent to P. R. Halmos, Department of Mathematics, University of California, Santa Barbara, CA 93106.

It is easy to be too busy to pay attention to what anyone else is doing, but not good. All of us should know, and want to know, what has been discovered since our formal education ended, but new words, and relations between them, are growing too fast to keep up. It is possible for a person to learn of the title of a recent work and of the key words used in it and still not have the faintest idea of what the subject is.

Progress Reports is to be an almost periodic column intended to increase everyone's mathematical information about what others have been up to. Each column will report one step forward in the mathematics of our time. The purpose is to inform, more than to instruct: what is the name of the subject, what are some of the words it uses, what is a typical question, what is the answer, who found it. The emphasis will be on concrete questions and answers (theorems), and not on general contexts and techniques (theories). References will be kept minimal: usually they will include only one of the earliest papers in which the answer appears and a more recent exposition of the discovery, whenever one is easily available.

Everyone is invited to nominate subjects to be reported on and authors to prepare the reports. The ground rules are that the principal theorem should be old enough to have been published in the usual sense of that word (and not just circulated by word of mouth or in preprints); it should be of interest to more than just a few specialists; and it should be new enough to have an effect on the mathematical life of the present and near future. In practice most reports will probably be on progress achieved somewhere between 5 and 15 years ago.

BERNOULLI SHIFTS

P. R. HALMOS

Many of the important problems of mathematics are of this kind: when are two objects the "same"? Example: when are two matrices similar?, when are two groups isomorphic?, when are two topological spaces homeomorphic?, when are two continuous mappings homotopic?

Ergodic theory (or at least a large part of it) is the study of measure-preserving transformations, and one of the important problems of ergodic theory is to determine when two measure-preserving transformations are "conjugate". If, for instance, S is translation by one unit to the right on the real line ($Sx = x + 1$) and T is its inverse ($Tx = x - 1$), then S and T are conjugate elements of the group of all measure-preserving transformations of the line; the reflection $Q(Qx = -x)$ transforms S onto $T(Q^{-1}SQ = T)$. If, for another example, S is rotation by 1 radian on the perimeter of the unit circle and T is rotation by 2 radians, then it is perhaps intuitively plausible that S and T are *not* conjugate, and indeed they are not, but the proof needs slightly more sophisticated methods. For a final example, consider the same S (rotation by 1 radian), but, this time, let T be translation modulo 2π by one unit to the right on the interval $[0, 2\pi)$. In this last example S and T act on different spaces, but, plainly, the difference is merely notational; if Q is the mapping $x \to e^{ix}$ from $[0, 2\pi)$ to the unit circle, then $Q^{-1}SQ = T$. The point of this example is to emphasize that conjugacy is a relation between measure-preserving transformations acting on possibly different spaces.

A special case of the conjugacy problem that was unsolved for a long time concerned "Bernoulli shifts". The prototypical Bernoulli shift is the measure-theoretic model of coin tossing. The space is the set of all two-way infinite sequences of 0's and 1's (heads and tails, from the beginning of time to eternity). Measure is defined by the familiar requirements that the probabilities of heads and tails be 1/2 and 1/2, and that the tosses on different days be independent of one another. The transformation is the index-shift one unit to the right (so that, for instance, the history of the coin experiment that results in heads today but results in tails on all past and future days is mapped onto the history that registers heads tomorrow and tails all other days). A more general Bernoulli shift allows n outcomes ($n \geqq 2$), with possibly unequal probabilities $p_1, \ldots, p_n$, but is otherwise formally the same. (The most general Bernoulli shift allows certain infinite experiments too.) The conjugacy problem for Bernoulli shifts is to decide when the shift built on $(p_1, \ldots, p_n)$ is conjugate to the one built on $(q_1, \ldots, q_m)$.

The first step toward the solution was taken by Kolmogorov (1958) and Sinai (1959). It depends on an ingenious conjugacy invariant, called *entropy*, suggested by statistical mechanics and information theory.

The definition of entropy goes like this. Suppose that X is a space with measure m. Define the entropy of a finite partition $\{A_1, \ldots, A_m\}$ of X to be the number $-\sum_{i=1}^{m} m(A_i) \log m(A_i)$. If T is a

measure-preserving transformation on X, consider the partition obtained from the A_i's by adjoining to them the $T^{-1}A_j$'s. (Example: if X is the heads-tails sequence space, and T is the shift, and if the partition $\{A_1, A_2\}$ is the two-way classification "today it's heads" and "today it's tails", then the finer partition is "today it's heads and tomorrow it's heads", "today it's heads and tomorrow it's tails", etc.) Calculate the entropy of the refined partition, and divide by 2; apply T again, calculate the entropy, and divide by 3; etc. The limit exists (that's not hard to prove); it is called the entropy of T relative to the originally chosen partition. The entropy of T is the supremum of all these relative entropies (for all possible partitions). The entropy of a partition corresponds to the amount of information that can be gained by performing an experiment; the limiting entropy relative to a partition corresponds to the average information per day that can be gained by repeated performances of the experiment; the entropy of a transformation corresponds to the maximum amount of information per day that can be gained, from the given model, by performing a finite experiment.

It is plausible and provable that the entropy of the Bernoulli shift built on $(p_1, \ldots, p_n)$ is the same as the entropy of the easiest partition associated with the shift (that is, $-\sum_{i=1}^{m} p_i \log p_i$). This is the Kolmogorov–Sinai solution of one of the unsolved problems of conjugacy: it implies that some very different looking Bernoulli shifts (e.g., the one for a coin and the one for a die) are indeed different. In other words: some Bernoulli shifts are *not* conjugate.

Is entropy a complete conjugacy invariant? If two transformations have the same entropy, are they necessarily conjugate? The answer is no. Next question: does the answer change to yes if the transformations to be considered are suitably and usefully restricted? For instance: is it true that if two Bernoulli shifts have the same entropy, then they are conjugate? The answer is yes. This is the first and greatest step in a sequence of results obtained by Ornstein (1970). It answered an old question, and, apparently, encouraged many scholars to attack the subject with new vigor.

Reference

D. Ornstein, Bernoulli shifts with the same entropy are isomorphic, Advances in Math., 4 (1970) 337–352 (MR 41 # 1973).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, SANTA BARBARA, CA 93106.

MATHEMATICAL NOTES

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Department of Mathematics, University of Wisconsin, Madison, WI 53706.

ON A CLASS OF ORDINARY LINEAR DIFFERENTIAL EQUATIONS HAVING $\sum_{k=0}^{\infty} C_k x^k$ AND $\sum_{k=0}^{n} C_k x^k$ AS SOLUTIONS

DAVID ZEITLIN

1. Introduction. Hoyt [1] has noted that the differential equation

$$xy'' - (x+n)y' + ny = 0, \qquad (n = 1, 2, 3, \ldots), \tag{1.1}$$

has the two solutions $y = e^x = \sum_{k=0}^{\infty} x^k/k!$ and its partial sum $y_n(x) = \sum_{k=0}^{n} x^k/k!$. Leighton [4] has

also discussed (1.1). In terms of differential operators $D \equiv d/dx$, we obtain from (1.1)

$$(xD - n)(D - 1)y = 0, \qquad (n = 1,2,3,\ldots). \tag{1.2}$$

Recently, Newton [5] considered differential equations of the form

$$(xD - n)(P(x)D^2 + Q(x)D + R(x))y = 0, \tag{1.3}$$

where P, Q, and R are analytic at $x = 0$. When P, Q and R are constants, his results for the same example can be compared to our Theorem 1 for $p = 2$ (see Section 2). A general result, having (1.2) as a special case, is given in Section 2.

2. The main result. Our general result is as follows:

THEOREM 1. *Let $y(x)$ be a solution of*

$$(D^p + A_{p-1}D^{p-1} + A_{p-2}D^{p-2} + \cdots + A_1D + A_0)y = 0, \qquad (p = 1,2,\ldots), \tag{2.1}$$

where A_i, $i = 0,1,2,\ldots,p$, are constants, with $A_0 \neq 0$, and $A_p = 1$. Then $y(x) = \Sigma_{k=0}^{\infty} C_k x^k$ and its partial sum $y_n(x) = \Sigma_{k=0}^{n} C_k x^k$ are both solutions of

$$\left[\prod_{j=0}^{p-1} (xD - (n - j))\right] \cdot \left(\sum_{i=0}^{p} A_i D^i\right) y = 0, \qquad (n \geqq p). \tag{2.2}$$

Proof. Since $y(x)$ is a solution of (2.1), it is also a solution of (2.2). By (2.1), we get

$$\sum_{j=0}^{p} A_j D^j \left(\sum_{k=0}^{\infty} C_k x^k\right) = \sum_{j=0}^{p} A_j \sum_{k=j}^{\infty} j! \binom{k}{j} C_k x^{k-j} = \sum_{j=0}^{p} j!\, A_j \sum_{i=0}^{\infty} \binom{i+j}{j} C_{i+j} x^i \equiv 0.$$

Equating coefficients of x^k to zero, we obtain the identity

$$\sum_{j=0}^{p} j! \binom{k+j}{j} A_j C_{k+j} = 0, \qquad (k = 0,1,2,\ldots). \tag{2.3}$$

Let $H(j,k)$ denote the summand in (2.3). If the polynomial operator of (2.1) is now applied to $y_n(x)$, the partial sum of $y(x)$, we obtain in a similar manner that

$$\sum_{j=0}^{p} A_j D^j \left(\sum_{i=0}^{n} C_i x^i\right) = \sum_{j=0}^{p} A_j \sum_{i=j}^{n} j! \binom{i}{j} C_i x^{i-j} = \sum_{j=0}^{p} \sum_{k=0}^{n-j} H(j,k) x^k.$$

If we apply the following identity,

$$\sum_{j=0}^{p} \sum_{k=0}^{n-j} H(j,k) x^k = \sum_{k=0}^{n-p} \left(\sum_{j=0}^{p} H(j,k)\right) x^k + \sum_{k=n-p+1}^{n} \left(\sum_{j=0}^{n-k} H(j,k)\right) x^k, \tag{2.4}$$

we obtain, noting that $\Sigma_{j=0}^{p} H(j,k) = 0$ by (2.3),

$$\sum_{j=0}^{p} A_j D^j y_n(x) = \sum_{k=n-p+1}^{n} \left(\sum_{j=0}^{n-k} j! \binom{k+j}{j} A_j C_{k+j}\right) x^k, \qquad (n \geqq p). \tag{2.5}$$

Since $(xD - (n-j))(R_n x^{n-j}) \equiv 0$, where R_n is a coefficient independent of x, it follows that $\Pi_{j=0}^{p-1}(xD - (n-j))$ annihilates all powers of x on the right hand side of (2.5), and thus $y_n(x)$ is also a solution of (2.2).

REMARKS. For $p = 1$, (2.2), with $A_1 = 1$ and $A_0 = a \neq 0$, gives $(xD - n)(D + a)y = 0$, which simplifies to $xy'' + (ax - n)y' - any = 0$, whose solutions are $y = e^{-ax}$ and $y_n(x) = \Sigma_{k=0}^{n}(-ax)^k/k!$. For $a = -1$, we obtain (1.1).

For $p = 2$, consider $(D^2 + 1)y = 0$, an example cited in [5, p. 595, (12)], wherein $y = \sin x$ and $y_{2n+1}(x) = \Sigma_{k=0}^{n}(-1)^k x^{2k+1}/(2k+1)!$ are shown to be solutions of $(xD - (2n+1))(D^2 + 1)y = 0$. Although $y = \sin x + \cos x = \Sigma_{k=0}^{\infty} C_k x^k$, where $C_k \neq 0$, $k = 0,1,\ldots$, is also a solution of $(D^2 + 1)y = 0$,

there is no $n \geqq 1$ such that $y_n(x) = \Sigma_{k=0}^n C_k x^k$ is a solution of $(xD - n)(D^2 + 1)y = 0$ (see (1.3)). An extension to $p \geqq 3$ of Newton's criterion [5, p. 595, top line], which depends essentially on the even or odd nature of $y(x)$, appears to be impractical, if not impossible. For comparison, our result (2.2) states that $y = \sin x + \cos x$ and its partial sum $y_n(x)$ are solutions of $(xD - (n-1))(xD - n)(D^2 + 1)y = 0$, which is also satisfied by

$$y = \sin x \quad \text{and} \quad y_n(x) = \sum_{k=0}^{[(n-1)/2]} (-1)^k x^{2k+1}/(2k+1)!,$$

as well as by $y = \cos x$ and its corresponding $y_n(x)$.

The following result is an application of (2.5) for $p = 1$.

LEMMA. *Let*

$$(D + A_0)y = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0, \tag{2.6}$$

where $A_0 \neq 0$ and a_i, $i = 0, 1, \ldots, m$, are constants. Then a particular integral $Z_p(x)$ of (2.6) is given by

$$Z_p(x) = \sum_{k=0}^m \left(\frac{(-1)^k}{k!} \sum_{n=k}^m n!(-1)^n a_n A_0^{k-n-1} \right) x^k. \tag{2.7}$$

Proof. If $y_n(x)$ is the partial sum of $y(x) = e^{-A_0 x} = \Sigma_{k=0}^\infty C_k x^k$, where $(D + A_0)y(x) = 0$, then by (2.5), for $p = 1$, $\Sigma_{j=0}^1 A_j D^j y_n(x) = A_0 C_n x^n$. Thus, if $a_n x^n$ is a term on the right hand of (2.6), then $b_n y_n(x)$ is a particular integral provided $b_n A_0 C_n x^n \equiv a_n x^n$, which holds if we choose $b_n = a_n/(A_0 C_n)$. Since particular integrals are additive, a particular integral for the polynomial in (2.6) is given by

$$Z_p(x) = \sum_{n=0}^m b_n y_n(x) = \sum_{n=0}^m b_n \sum_{k=0}^n C_k x^k = \sum_{k=0}^m C_k x^k \sum_{n=k}^m (a_n)/(A_0 C_n),$$

which yields (2.7), since $C_k = (-A_0)^k/(k!)$, $k = 0, 1, \ldots$.

3. A generalization of (1.1).

THEOREM 2. *Let $P(x)$, with $P'(x) \not\equiv 0$, be analytic at $x = 0$. Let $y(x)$ be a solution of $(D - P'(x))y = 0$. Then $y(x) = e^{P(x)} = \Sigma_{k=0}^\infty P^k(x)/k!$ and $Y_n(x) = \Sigma_{k=0}^n P^k(x)/k!$ are both solutions of*

$$\left(D - \frac{P''(x)}{P'(x)} - n\frac{P'(x)}{P(x)} \right)(D - P'(x))y = 0. \tag{3.1}$$

Proof. Since $y(x)$ is again a solution of (3.1), we consider now

$$\begin{aligned} (D - P'(x))Y_n(x) &= \sum_{k=1}^n kP^{k-1}(x)P'(x)/k! - P'(x)\sum_{k=0}^n P^k(x)/k! \\ &= P'(x)\left(\sum_{j=0}^{n-1} P^j(x)/j! - \sum_{k=0}^n P^k(x)/k! \right) = -P'(x)P^n(x)/n! \\ &= -D(P^{n+1}(x))/(n+1)!. \end{aligned}$$

Thus, if $(D - R(x))D(P^{n+1}(x)) \equiv 0$, then $R(x) \equiv D^2(P^{n+1}(x))/D(P^{n+1}(x))$. Since

$$D(P^{n+1}(x)) = (n+1)P^n(x)P'(x) \quad \text{and} \quad D^2(P^{n+1}(x)) = (n+1)(P^n(x)P''(x) + nP^{n-1}(x)(P'(x))^2),$$

we obtain $R(x)$ as given in (3.1), thus assuring that $Y_n(x)$ is also a solution of (3.1).

REMARKS. If we set $P(x) = -ax$, then (3.1) gives $(D - (n/x))(D + a)y = 0$, which gives (1.1) for $a = -1$. For applications, choices of $P(x)$ in (3.1) can be polynomials, e^{bx}, $x^r e^{bx}, \ldots$. For some

choices of $P(x)$, the second solution $Y_n(x)$ is not a true partial sum $y_n(x)$ of $y(x)$. In Reddick [**6**, pp. 256–257, example 3], (1.1), with $n=2$, is shown to have the two solutions e^x and $y_2(x)$. In [**6**, pp. 242–243], example 1 is (3.1) for $P=-2x$ and $n=1$. In [**6**, p. 245], two special cases of (3.1) for $n=1$ occur, where $P(x)=2x$, x^2, respectively, in problems 3,5. In [**3**, p. 427, (2.111)], the two solutions of (1.1) for $n=1$ are given.

For $P(x)=x^2$, we obtain from (3.1)

$$x^2y''-(2x^3+(2n+1)x)y'+4nx^2y=0, \tag{3.2}$$

whose solutions are e^{x^2} and $Y_n(x)=\Sigma_{k=0}^n x^{2k}/k!$. In Kamke [**3**, p. 451, (2.215)], the differential equation cited is of the form

$$x^2y''+(ax^{m+1}+bx)y'+(Ax^{2m}+Bx^m+C)y=0. \tag{3.3}$$

For $a=-2$, $m=2$, $b=-(2n+1)$, $A=0$, $B=4n$, and $C=0$, (3.3) gives (3.2). In [**3**, p. 450, (2.210)], the differential equation cited is of the form

$$x^2y''-(2x^3-2ax)y'+(2mx^2+((-1)^m-1)a)y=0. \tag{3.4}$$

For $m=2n$ and $-2a=2n+1$, (3.4) gives (3.2). We note that (3.4) also appears in Szegö [**8**, p. 377, problem 25].

In the notation of Theorem 2, one easily shows that $y(x)=P^a(x)e^{P(x)}$ and $P^a(x)Y_n(x)=\Sigma_{k=0}^n P^{a+k}(x)/k!$ are both solutions of the differential equation

$$\left(D-\frac{P''(x)}{P'(x)}-(n+a)\frac{P'(x)}{P(x)}\right)\left(D-P'(x)-a\frac{P'(x)}{P(x)}\right)y=0. \tag{3.5}$$

For $a=0$, (3.5) gives (3.1). We note that (3.5) can also be obtained from (1.1) by a change of dependent variable defined by $Y(x)=P^a(x)y(P(x))$.

For $p=2$, (2.2), with $A_2=1$, gives $\Sigma_{k=0}^4 R_k(x)D^ky=0$, where $R_0(x)=n(n-1)A_0$,

$$R_1(x)=2(1-n)A_0x+n(n-1)A_1,\quad R_2(x)=A_0x^2-2(n-1)A_1x+n(n-1),$$
$$R_3(x)=A_1x^2-2(n-1)x,$$

and $R_4(x)=x^2$. Omitting details, one easily shows that (2.2) can be simplified to read as

$$\sum_{k=0}^{2p}\left\{\sum_{i=0}^{k}\left(A_{k-i}\sum_{j=i}^{p}\mathfrak{S}_j^i\sum_{r=j}^{p}\binom{r}{j}(p-1-n)^{r-j}S_p^r\right)x^i\right\}D^ky=0,\qquad (n\geqq p), \tag{3.6}$$

where S_p^r and $\mathfrak{S}_j^i$ are Stirling numbers of the first and second kind, respectively; see Jordan [**2**, Chapter 4]. For $k=0$, the coefficient of y in (3.6) is $A_0\Pi_{j=0}^{p-1}(p-1-j-n)$. Thus, if $n=p-1-j$, $j=0,1,\dots,p-1$, the order of (3.6) is reducible by one.

It should be noted that (1.1) can be obtained from the differential equation satisfied by the Laguerre polynomials $L_n^{(a)}(x)$. Recall (see [**8**, p. 99, (5.1.2); p. 102, (5.3.5)]) that

$$xy''+(a+1-x)y'+ny=0 \tag{3.7}$$

is satisfied by the two independent solutions

$$y_1^*(x)=L_n^{(a)}(x)=\sum_{k=0}^{n}\binom{n+a}{n-k}(-x)^k/k!\quad\text{and}\quad y_2^*(x)=x^{-a}\sum_{k=0}^{\infty}\left[\binom{n+a}{k}\Big/\binom{a-1}{k}\right]x^k/k!.$$

If we set $a=-1-n$ in (3.7), we obtain (1.1). Since

$$\binom{-1}{n-k}=(-1)^{n-k}\quad\text{and}\quad\binom{-n-2}{k}=(-1)^k\binom{n+1+k}{k},$$

we see that $y_1^*(x)=(-1)^n\,\Sigma_{k=0}^n x^k/k!=(-1)^ny_n(x)$ and

$$y_2^*(x) = (n+1)! \sum_{k=0}^{\infty} x^{n+1+k}/(n+1+k)! = (n+1)! \sum_{j=n+1}^{\infty} x^j/j! = (n+1)! \left(e^x - \sum_{j=0}^{n} x^j/j!\right)$$
$$= (n+1)!\, e^x - (n+1)!\, y_n(x).$$

Thus, we have $y(x) = e^x$ and $y_n(x)$ as solutions of (1.1).

References

1. J. P. Hoyt, Remarks on a second order differential equation, this MONTHLY, 74 (1967) 436.
2. C. Jordan, Calculus of Finite Differences, Chelsea, New York, 1960.
3. E. Kamke, Differentialgleichungen, Lösungsmethoden und Lösungen, Vol. 1, 2nd edition, Akademische Verlagsgesellschaft, Leipzig, 1943.
4. W. Leighton, On the differential equation $xy'' - (x+n)y' + ny = 0$, this MONTHLY, 76 (1969) 52–53.
5. T. A. Newton, On using a differential equation to generate polynomials, this MONTHLY, 81 (1974) 592–601.
6. H W. Reddick, Differential Equations, Wiley, New York, 1950.
7. I. J. Schwatt, An Introduction to the Operations with Series, 2nd edition, Chelsea, New York, 1962, pp. 181–182.
8. G. Szegö, Orthogonal Polynomials, Amer. Math. Soc., New York, 1959.

1650 VINCENT AVENUE NORTH, MINNEAPOLIS, MN 55411.

COMMUTATORS AND THE COMMUTATOR SUBGROUP

I. M. ISAACS

The commutator subgroup G' of a group G is generated by commutators, elements of the form $[x, y] = x^{-1}y^{-1}xy$. As is quite well known, not every element of G' need be a commutator. What is perhaps less well known is a convenient source of finite groups which are examples of this phenomenon. The purpose of this note is to provide such a source. (Other examples are described in **[1]**, **[2]** and **[3]**.) The method given here can be used to construct both solvable and nonsolvable groups and even yields examples which are perfect, that is $G' = G$. The author is unaware, however, of any nonabelian *simple* group which contains a noncommutator.

Our examples will be wreath products and we begin with a description of these groups. Let U and H be any groups. Their wreath product $G = U \wr H$ has as normal subgroup the group B of all functions $f: H \to U$. Multiplication in B is pointwise. Also $H \subseteq G$ and $G = BH$ (and of course $B \cap H = 1$). Finally, to complete the description of G we have for $f \in B$ and $h \in H$ that $h^{-1}fh = f^h \in B$ with $f^h(x) = f(xh^{-1})$ for $x \in H$. We refer to B as the *base group* of the wreath product.

THEOREM. *Let U and H be finite groups with U abelian and H nonabelian. Let $G = U \wr H$. Then G' contains a noncommutator if*

$$(*) \qquad \sum_{A \in \mathscr{A}} \left(\frac{1}{|U|}\right)^{|H:A|} \leqq \frac{1}{|U|},$$

where $\mathscr{A}$ is the set of maximal abelian subgroups of H. In particular, this condition holds whenever $|U| \geqq |\mathscr{A}|$.

Actually, as can be seen from the proof, somewhat weaker conditions that (*) suffice although they are hard to state cleanly. In fact, if H is nonabelian of order 6, we can take $|U| = 2$. Although (*) is not satisfied, nevertheless the resulting group G of order $2^7 \cdot 3$ is an example where G' contains a noncommutator.

LEMMA 1. *Let G be a group with abelian $A \lhd G$ and suppose $G = AH$ with $A \cap H = 1$. If $[x, y] \in A$ with $x, y \in G$, then $[x, y] \in [A, K]$ for some abelian $K \subseteq H$.*

Proof. Write $x = ah$ and $y = bk$ with $a, b \in A$ and $h, k \in H$. Since $[x, y] \in A$, the images of x and y in G/A commute. Since these are also the images of h and k, it follows that $[h, k] \in A$. Since also $[h, k] \in H$, we have $[h, k] = 1$ and $K = \langle h, k\rangle$ is abelian.

Now $[A, K]$, the group generated by commutators of elements of A with elements of K, is normal in AK. The images of A and K in $AK/[A, K]$ are abelian and centralize each other, and hence $AK/[A, K]$ is abelian. Thus $[x, y] \in (AK)' \subseteq [A, K]$. ∎

LEMMA 2. *Let $G = U \wr H$ where U is abelian and G is finite. Let B be the base group of G and let $K \subseteq H$. Then $|[B, K]| = |U|^{|H|-|H:K|}$.*

Proof. Let T be a set of representatives for the left cosets of K in H. For each $t \in T$, define $\sigma_t : B \to U$ by $\sigma_t(f) = \Pi_{k \in K} f(tk)$. Then σ_t is a homomorphism and $\sigma_t(f^k) = \sigma_t(f)$ for $f \in B$ and $k \in K$.

Let $C = \bigcap_{t \in T} \ker \sigma_t$. Then $|C| = |U|^{|H|-|H:K|}$ since any f in C may be specified arbitrarily on all but one element in each coset. We claim that $[B, K] = C$.

To show that $[B, K] \subseteq C$, let $f \in B$. Then $[f, k] = f^{-1}f^k$ and $\sigma_t([f, k]) = \sigma_t(f^{-1})\sigma_t(f^k) = 1$. Thus $[f, k] \in C$ and hence $[B, K]$, the group generated by all $[f, k]$, is contained in C.

To show that $C \subseteq [B, K]$, let $\tau : B \to B/[B, K]$ be the canonical homomorphism and note that $\tau(f^k) = \tau(f)$ for $f \in B$. Let $c \in C$ and $k \in K$ and define $c_k \in B$ by

$$c_k(x) = \begin{cases} c(x) & \text{if } xk \in T \\ 1 & \text{if } xk \notin T. \end{cases}$$

It follows that $c = \Pi_{k \in K} c_k$. Let $b = \Pi_{k \in K} (c_k)^k$. Since $\tau(f) = \tau(f^k)$ we have $\tau(c) = \tau(b)$. We claim that $b = 1$ and thus $\tau(c) = 1$ and $C \subseteq [B, K]$. We compute $b(x)$ for $x \in H$. If $x \notin T$, then $(c_k)^k(x) = c_k(xk^{-1}) = 1$ for all k and $b(x) = 1$. If $x \in T$, then

$$(c_k)^k(x) = c_k(xk^{-1}) = c(xk^{-1})$$

and so $b(x) = \Pi_k c(xk^{-1}) = \sigma_x(c) = 1$. The proof is complete. ∎

Proof of Theorem. Let B be the base group of $G = U \wr H$. Then $[B, H] \subseteq G'$ and $|[B, H]| = |U|^{|H|-1}$ by Lemma 2. If every element of $[B, H]$ is a commutator, then $[B, H] = \bigcup_{A \in \mathcal{A}} [B, A]$ by Lemma 1. Since $|[B, A]| = |U|^{|H|-|H:A|}$, this forces

$$\sum_{A \in \mathcal{A}} |U|^{|H|-|H:A|} > |U|^{|H|-1}$$

and thus

$$\sum_{A \in \mathcal{A}} \left(\frac{1}{|U|}\right)^{|H:A|} > \left(\frac{1}{|U|}\right)$$

and the first statement is proved. The second statement follows since $|H:A| \geqq 2$ for all $A \in \mathcal{A}$. ∎

We remark that if H is simple, U is abelian and $G = U \wr H$, then $G' = G''$. Thus if U is large enough, then G' is a perfect group in which not every element is a commutator.

Finally we mention that one can read off from the character table of a group, the elements which are commutators. In fact $g \in G$ is a commutator iff

$$\sum \chi(g)/\chi(1) > 0,$$

where the sum runs over all complex irreducible characters χ of G.

References

1. R. D. Carmichael, Groups of finite order, Dover, New York, 1937.
2. I. D. MacDonald, On cyclic commutator subgroups, J. London Math. Soc., 38 (1963) 419–422.
3. W. R. Scott, Group theory, Prentice-Hall, Englewood Cliffs, N. J., 1964.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN, MADISON, WI 53706.

RESTRICTIONS ON THE VALUES OF DERIVATIVES

WALTER RUDIN

In **[1]**, F. D. Hammer asked whether there exists a differentiable function f with $f(r)$ rational but $f'(r)$ irrational for every rational r. Posed this way, the problem involves arithmetic properties of the real numbers, and the explicit example constructed by W. Knight **[2]** makes full use of these arithmetic features.

However, the phenomenon under consideration really depends only on the fact that the set of all rational numbers is countable and dense in the line, and that the irrationals are also dense. (Another solution of the problem, found by Dan Simchoni and stated without proof after **[2]**, furnishes an entire function with restrictions of f and f' on an arbitrary countable set, and makes no use of arithmetic.) Once this is recognized, it is easy (as we shall see) to extend this phenomenon to infinitely differentiable functions, in any finite number of variables.

Let n be a fixed positive integer. A *multi-index* is an ordered n-tuple $\alpha = (\alpha_1, \ldots, \alpha_n)$ in which each α_i is a nonnegative integer. To each multi-index α corresponds a differential operator

$$D^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n}.$$

As usual, R is the real line, R^n is euclidean n-space, and $C^\infty(R^n)$ is the class of all functions $f: R^n \to R$ with $D^\alpha f$ continuous for every α.

THEOREM. *Suppose that*
(a) *A is a countable subset of R^n, and*
(b) *for each multi-index α, B_α is a dense subset of R.*
Then there exists an $f \in C^\infty(R^n)$ such that $D^\alpha f$ maps A into B_α, for every α.

Proof. We shall use the customary multi-index notations

$$|\alpha| = \alpha_1 + \cdots + \alpha_n, \quad \alpha! = \alpha_1! \cdots \alpha_n!, \quad x^\alpha = \xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}$$

if $x = (\xi_1, \ldots, \xi_n) \in R^n$.

Arrange the members of A in a sequence $\{x_i\}$, $i = 0, 1, 2, \ldots$, with $x_i \neq x_j$ if $i \neq j$. For $i \geqq 0$, choose $\psi_i \in C^\infty(R^n)$ with compact support K_i, such that
(i) K_i contains no x_m with $m < i$,
(ii) $0 \leqq \psi_i(x) \leqq 1$ for all $x \in R^n$, and
(iii) $\psi_i(x) = 1$ for all x in some neighborhood of x_i.

Choose $c_0(\alpha) \in B_\alpha$, so small that the power series

$$g_0(x) = \sum_\alpha \frac{c_0(\alpha)}{\alpha!}(x - x_0)^\alpha \tag{1}$$

defines an entire function g_0, with $|g_0(x)| < 1$ on K_0. If $f_0 = \psi_0 g_0$ then $f_0 \in C^\infty(R^n)$, and

$$(D^\alpha f_0)(x_0) = (D^\alpha g_0)(x_0) = c_0(\alpha) \tag{2}$$

for every α, since $f_0 = g_0$ in a neighborhood of x_0.

Suppose that $m \geqq 1$, and that $f_i \in C^\infty(R^n)$ has been constructed for $0 \leqq i < m$, with support in K_i. Then there exist $c_m(\alpha) \in R$, so small that the function

$$g_m(x) = \sum_\alpha \frac{c_m(\alpha)}{\alpha!}(x - x_m)^\alpha \tag{3}$$

is entire, such that

$$|D^\beta(g_m\psi_m)(x)| < 2^{-m} \qquad (x \in R^n) \tag{4}$$

for all β with $|\beta| \leqq m$, and such that

$$c_m(\alpha) + \sum_{j=0}^{m-1} (D^\alpha f_j)(x_m) \in B_\alpha \tag{5}$$

for all α.

(Note that (4) can be achieved because only finitely many derivatives are involved and because ψ_m has compact support, and that (5) is possible because B_α is dense in R.)

Put $f_m = g_m\psi_m$, and proceed with the construction.

Define $f = \Sigma_0^\infty f_m$. By (4), $|D^\beta f_m| < 2^{-m}$ when $m \geqq |\beta|$. Therefore the series $\Sigma D^\beta f_m$ converges uniformly, for every β; it follows that $f \in C^\infty(R^n)$. If $j > m$ then f_j vanishes in a neighborhood of x_m. Hence $(D^\alpha f_j)(x_m) = 0$. Consequently, for every α, (5) gives

$$(D^\alpha f)(x_m) = D^\alpha(f_0 + \cdots + f_m)(x_m) \in B_\alpha \tag{6}$$

for every $x_m \in A$.

This work was partially supported by NSF Grant MPS 75–06687.

References

1. F. D. Hammer, Problem 5955, this MONTHLY, 81 (1974) 176.
2. William Knight, Solution of Problem 5955, this MONTHLY, 82 (1975) 415–416.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN-MADISON, MADISON, WI 53706.

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics and Statistics, The University of Calgary, Calgary, Alberta, Canada, T2N 1N4.

CAN ONE SOLVE EQUATIONS IN GROUPS?

JAN MYCIELSKI

We will discuss, for some classical groups G, the problem of solvability of equations of the form

$$x^{p_1}y^{q_1}x^{p_2}y^{q_2}\cdots x^{p_m}y^{q_m} = a, \tag{1}$$

where x, y are the unknowns to be found in G, $p_1, \ldots, p_m, q_1, \ldots, q_m$ are given integers and $a \in G$.

Let $\mathbf{Z}$, $\mathbf{R}$ and $\mathbf{C}$ denote as usual the rings of integers, reals, and complex numbers respectively. Let $SO(n)$ denote the group of all orthogonal matrices of size $n \times n$ with real entries and determinant 1,

$SL(n, A)$ the group of all matrices of size $n \times n$ with entries from A and determinant 1, where A is any commutative ring with unity, and $SU(n)$ the subgroup of unitary matrices of $SL(n, \mathbf{C})$. Then $SU(2)$ has a natural isomorphism to the multiplicative group of quaternions of norm 1. Let I be the unit matrix in the appropriate group; m denotes any positive integer and $p_1, \ldots, p_m, q_1, \ldots, q_m$ denote integers, all different from 0 except perhaps q_m.

PROBLEMS. (a) Let n be an even positive integer. Does the equation

$$x^{p_1}y^{q_1}\cdots x^{p_m}y^{q_m} = -I \tag{2}$$

have solutions x, y in $SL(n, \mathbf{R})$, or at least in $SL(n, \mathbf{C})$, for every $m, p_1, \ldots, p_m, q_1, \ldots, q_m$?

(b) Does the equation

$$x^{p_1}y^{q_1}\cdots x^{p_m}y^{q_m} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3}$$

have solutions x, y in $SO(3)$, or at least in $SL(3, \mathbf{C})$, for every $m, p_1, \ldots, p_m, q_1, \ldots, q_m$?

These problems are open and seem to be the most relevant cases of the corresponding more general problems, where the right sides are unspecified. We have the following remarks.

(i) In the case when $\Sigma p_i \neq 0$ or $\Sigma q_i \neq 0$ the answer is yes because all the groups $SO(n)$ are divisible. In fact every connected compact group G is divisible, i.e., each equation $x^m = a$ with $a \in G$ has a solution $x \in G$, see [7].

(ii) For every $m, p_1, \ldots, p_m, q_1, \ldots, q_m$ there exists a number $\alpha \geqq 0$ such that

$$\{x^{p_1}y^{q_1}\cdots x^{p_m}y^{q_m} : x, y \in SO(3)\} = \{z \in SO(3) : z \text{ rotates } \mathbf{R}^3 \text{ through an angle } \leqq \alpha\}.$$

This follows from the fact that $SO(3)$ and hence $SO(3) \times SO(3)$ are connected and the range of a continuous function over a connected space is connected and from the fact that inner automorphisms of $SO(3)$ rotate the axes of rotation in every possible way without changing the angles of rotation. Also inner automorphisms permit one to assume without loss of generality that $q_m \neq 0$.

(iii) Moreover $\alpha > 0$. This follows from a theorem of Hausdorff (later much generalized [**1, 2, 3, 5, 9**]) which tells us that $SO(3)$ has a nonabelian free subgroup. Problem (b) for $SO(3)$ is equivalent to the question if $\alpha = \pi$.

(iv) If the answer to (b) for $SO(3)$ is yes, then every equation of the form (1) with $a \in SO(3)$ has solutions x, y in $SO(3)$. This follows from (ii) since

$$a_0 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

rotates $\mathbf{R}^3$ through the angle π.

(v) A positive answer to (b) for $SO(3)$ is equivalent to the statement that for every sequence $m, p_1, \ldots, p_{2m}$ the equation

$$\operatorname{trace}(x^{p_1}a_0x^{p_2}a_0\cdots x^{p_{2m}}a_0) = -1,$$

where a_0 is as in (iv), has a solution $x \in SO(3)$. This can be proved by substituting yxy^{-1} for y in (3), a geometric argument similar to the one in (ii) and the fact that $\operatorname{trace}(z) = 1 + 2\cos(\text{rotation angle of } z)$ for every $z \in SO(3)$.

(vi) A positive answer to (b) for $SO(3)$ is equivalent to the statement that every equation of the form

$$\text{real part}\,(x^{p_1}\sqrt{-1}\,x^{p_2}\sqrt{-1}\cdots x^{p_{2m}}\sqrt{-1}) = 0$$

is satisfied by some quaternion $x \neq 0$. This is a simple translation of (v) into the language of quaternions. Also, the equation (3) can be solved in $SO(3)$ just if the equation

$$\text{real part}\,(x^{p_1}y^{q_1}\cdots x^{p_m}y^{q_m}) = 0$$

can be solved in quaternions $x \neq 0$ and $y \neq 0$.

(vii) To answer the problem (a) it would be enough to get an answer "yes" for $n = 2$. In fact for every even $n > 2$ a solution x, y of (2) could be constructed by juxtaposing solutions of size 2×2 along the main diagonals of x and y.

(viii) Similar questions may be considered for other groups. But easy calculations show that already

$$x^2 = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$$

has no solution in $SL(2, \mathbf{C})$. (Are there such examples for $SL(m, \mathbf{C})$ modulo its center ?) Also there are no quaternions $x \neq 0$ and $y \neq 0$ satisfying

$$[x^2, yxy^{-1}] = -1,$$

where $[u, v] = uvu^{-1}v^{-1}$. (This example is due to J. Browkin, see [**10**].) In fact elementary calculations show that

$$\text{real part}\,[x^2, yxy^{-1}] \geqq -5/27,$$

and this inequality is sharp. In the same vein, the equation

$$[x^2, yxy^{-1}] = -I$$

cannot be solved in $SO(4)$. In fact, the rotations of $\mathbf{R}^4$ can be written in the form

$$p \to apb^{-1},$$

where $p \in \mathbf{R}^4$ is treated as a quaternion and a and b are quaternions of norm 1. Then $-I$ is represented by $p \to -p$ and we must have $a = \pm 1$ and $b = \mp 1$ respectively. Hence, as told above, either a or b cannot be of the form $[x^2, yxy^{-1}]$. Therefore, $-I$ also is not of this form. Are there such examples for $SO(6)$, $SO(8), \ldots$ (or for $SU(4)$, $SU(6), \ldots$)?

(ix) T. Dekker [**2**] has proved that there exist x_0, y_0 in $SO(4)$ such that 1 is not an eigenvalue of

$$x_0^{p_1}y_0^{q_1}\cdots x_0^{p_m}y_0^{q_m},$$

for every $m, p_1, \ldots, p_m, q_1, \ldots, q_m$; and the same follows for $SO(4k)$, $k = 1, 2, \ldots$ (again by composing cells along the main diagonals). However for $SO(6)$ this problem is open. A positive answer would imply the existence of some remarkable partitions of the spheres S^5, S^9, $S^{13}, \ldots$ which are known to exist for the spheres S^3, S^7, $S^{11}, \ldots$ and the spaces $\mathbf{R}^m$ with $m \geqq 3$. E.g.: Partitions into three parts A, B and C such that

$$A \equiv B \equiv C \equiv A \cup B \equiv B \cup C \equiv A \cup C,$$

where $X \equiv Y$ means that there exists a rigid and orientation preserving motion r such that $r(X) = Y$, see [**2, 4, 8, 11**].

(x) Let us add that W. Magnus and B. H. Neumann [**6, 12**] have proved the existence of two matrices x_0 and y_0 in $SL(2, \mathbf{Z})$, e.g.,

$$x_0 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \quad \text{and} \quad y_0 = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix},$$

such that for all $m, p_1, \ldots, p_m, q_1, \ldots, q_m$ we have

$$\text{trace}(x_0^{p_1}y_0^{q_1}\cdots x_0^{p_m}y_0^{q_m}) \neq 2.$$

This inequality is equivalent to the statement that 1 is not an eigenvalue of $x_0^{p_1}y_0^{q_1}\cdots x_0^{p_m}y_0^{q_m}$. Using the results in [**2, 5, 9, 11**] this allows us to construct curious partitions of the kind mentioned in (ix) of the spaces $\mathbf{C}^2-\{(0,0)\}$, $\mathbf{R}^2-\{(0,0)\}$ and $\mathbf{Z}^2-\{(0,0)\}$, with $r \in SL(2,\mathbf{Z})$.

References

1. S. Balcerzyk and J. Mycielski, On faithful representations of free products of groups, Fund. Math., 50 (1961) 63–71.
2. T. J. Dekker, Decompositions of sets and spaces I, II and III, Indag. Math., 18 (1956) 581–595 and 19 (1957) 104–107.
3. ———, On reflections of Euclidean spaces generating free products, Nieuw Arch. Wisk., 7 (1959) 57–60.
4. ———, On free groups of motions without fixed points, Indag. Math., 20 (1958) 348–353.
5. K. Kuratowski, Applications of the Baire-category method to the problem of independent sets, Fund. Math., 81 (1973) 65–72.
6. W. Magnus, Rational representations of Fuchsian groups and nonparabolic subgroups of the modular group, Nachr. Akad. Wiss. Göttingen II, Mathem. -Phys. Klasse, Jahrgang 1973, No. 9, 179–189.
7. J. Mycielski, Some properties of connected compact groups, Coll. Math., 5 (1958) 162–166.
8. ———, On the paradox of the sphere, Fund. Math., 42 (1955) 348–355.
9. ———, Almost every function is independent, Fund. Math., 81 (1973) 43–48.
10. ———, Obroty i slowa, Wiadomośći Matem., 10 (1967) 3–5. (In Polish)
11. J. Mycielski and S. Świerczkowski, On free groups of motions and decompositions of Euclidean spaces, Fund. Math., 45 (1958) 283–291.
12. B. H. Neumann, Über ein gruppentheoretisch-arithmetisches Problem, Sitzungsber. Preuss. Akad. Wiss. Phys. Math. Klasse, No. X(1933), 18 pp.

MATHEMATICS DEPARTMENT, UNIVERSITY OF COLORADO, BOULDER, CO 80309.

CLASSROOM NOTES

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Department of Mathematics, University of Wisconsin, Madison, WI 53706.

ON HALLEY'S VARIATION OF NEWTON'S METHOD

GEORGE H. BROWN, JR.

1. Introduction. Newton's method, properly known as the Newton–Raphson method [**5**, pp. 575–78], iteratively approximates a *simple real* root α of the *real* equation $f(x)=0$. A sequence $\{x_n\}$ due to **Newton's method** is generated by the well-known recurrence relation

$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)} \qquad (n=0,1,2,\ldots). \tag{1.1}$$

Under sufficient conditions — such as those of Fourier stated by Hildebrand [**5**, p. 577] or the less restrictive ones of Henrici [**4**, p. 79] — a sequence $\{x_n\}$ computed by means of (1.1) converges to α.

Nearly three hundred years ago Halley [**3**], using the *second* derivative of f to accelerate the convergence of $\{x_n\}$, presented an interesting variation of Newton's method. Frame [**1, 2**] and later Wall [**7**] published in this MONTHLY more recent derivations thereof, using a geometric approach involving a parabola. It is the purpose here to rederive Halley's variation of Newton's method in an analytic way and to evaluate briefly its practicality.

2. Derivation of Halley's variation. In general, for an arbitrary function g we know that the product fg preserves the zeros of the given function f; therefore, we may solve a new equation $f(x)g(x)=0$ rather than simply $f(x)=0$. The crux of the present derivation is to obtain a *specific* function g such that $(fg)''$ will be **zero** at α. This "increased linearity" in the neighborhood of α should result in faster convergence of $\{x_n\}$.

On a suitable interval I containing α, it is sufficient to assume that f and g are C^2 and that $f', g \neq 0$. Noting that $f(\alpha)=0$ and imposing the crucial condition that $[f(\alpha)g(\alpha)]''=0$, we expand $(fg)''$ to obtain

$$\frac{g'(\alpha)}{g(\alpha)} = -\frac{f''(\alpha)}{2f'(\alpha)}. \tag{2.1}$$

Integration of (2.1) over *all* of I with appropriate choice of constant yields

$$g = (f')^{-1/2}.$$

The final step is to rewrite (1.1) in terms of the product fg rather than just f. A sequence $\{x_n\}$ due to **Halley's variation** is generated by the resulting recurrence relation

$$x_{n+1} = x_n - \frac{1}{[f'(x_n)/f(x_n)] - [f''(x_n)/2f'(x_n)]} \qquad (n = 0,1,2,\dots). \tag{2.2}$$

Under sufficient conditions — identical to those already cited for Newton's method — a sequence $\{x_n\}$ computed by means of (2.2) converges to α.

3. Convergence. A convergent iterative process is said to be of **order k** if $|\alpha - x_{n+1}|$ tends to be proportional to $|\alpha - x_n|^k$ as $n \to \infty$ [**5**, p. 578], a concept originated by Schröder [**6**]. It is known [**5**, pp. 575–76] that Newton's method is *second*-order ($k=2$); moreover, Frame [**2**] and Wall [**7**] have demonstrated that Halley's variation is *third*-order ($k=3$). Therefore, as anticipated, Halley's error-*cubing* variation produces faster convergence of $\{x_n\}$ than does Newton's error-*squaring* method.

4. Conclusion. A count of function evaluations and algebraic operations necessitated by (1.1) and (2.2) indicates that **two** iterations of Newton's method and **one** iteration of Halley's variation require roughly the same computing time. Furthermore, faster convergence of $\{x_n\}$ is attained by the *fourth*-order combination of **two** consecutive iterations of Newton's *second*-order method than by **one** iteration of Halley's *third*-order variation. Consequently — in terms of computing time and convergence rate — a **pair** of iterations of Newton's method is, in general, more efficient than a **single** iteration of Halley's variation. In view of this consideration, Halley's variation is *not* an improvement of Newton's method but merely an interesting modification thereof.

5. Example. The first positive root, $\alpha = 0.860\,333\,589$, of

$$x - \cot x = 0$$

may be initially approximated by $x_0 = \pi/4$. With Newton's method $|\alpha - x_1| = 0.003\,401\,480$ and $|\alpha - x_2| = 0.000\,006\,351$; with Halley's variation $|\alpha - x_1| = 0.000\,180\,758$.

G. H. Brown, J. W. Brown, and the referees made suggestions concerning this note.

References

1. J. S. Frame, A variation of Newton's method, this MONTHLY, 51 (1944) 36–38.

2. ———, Remarks on a variation of Newton's method, this MONTHLY, 52 (1945) 212–214.

3. E. Halley, Methodus nova, accurata & facilis inveniendi radices aequationum quarumcumque generaliter, sine praevia reductione, Philos. Trans. Roy. Soc. London, 18 (1694) 136–148.

4. P. Henrici, Elements of Numerical Analysis, Wiley, New York, 1964.

5. F. B. Hildebrand, Introduction to Numerical Analysis, 2d ed., McGraw-Hill, New York, 1974.
6. E. Schröder, Ueber unendlich viele Algorithmen zur Auflösung der Gleichungen, Math. Ann., 2 (1870) 317–365.
7. H. S. Wall, A modification of Newton's method, this MONTHLY, 55 (1948) 90–94.

117 HUNT DRIVE, PRINCETON, NJ 08540.

ON A CLASS OF SEQUENCES OF INTEGERS

D. SURYANARAYANA

Let $\sigma(n)$ denote the sum of all the positive divisors of n and $\phi(n)$ denote the Euler Totient function. Let $\psi(n)$ denote the Dedekind function (cf. **[3]**, p. 123; also cf. **[2]**) defined by

$$\psi(n)=\sum_{d\delta=n}\mu^2(d)\delta,$$

where $\mu(n)$ is the Möbius function. Let $F_n=2^{2^n}+1$ and $M_{p_n}=2^{p_n}-1$ respectively denote the sequences of Fermat and Mersenne numbers, where p_n denotes the nth prime, counting $p_1=2$, $p_2=3$, etc.

In 1954, P. Erdös **[4]** proposed a problem which is equivalent to the following:

$$\frac{\sigma(F_n)}{F_n}\to 1 \quad \text{as} \quad n\to\infty. \tag{1}$$

In 1955, while giving a solution of (1), Ranko Bojanić **[1]** noted that the following also hold:

$$\frac{\phi(F_n)}{F_n}\to 1,\ \frac{\sigma(M_{p_n})}{M_{p_n}}\to 1,\ \frac{\phi(M_{p_n})}{M_{p_n}}\to 1 \quad \text{as} \quad n\to\infty. \tag{2}$$

In this note we prove that the properties (1) and (2) are enjoyed not merely by the two sequences F_n and M_{p_n} alone but by a class of sequences of positive integers to which F_n and M_{p_n} belong. In fact, we prove the following:

THEOREM. *Let $\{a_n\}$ be an increasing sequence of positive integers such that* $\log a_n/q_n \log q_n \to 0$ *as $n\to\infty$, where q_n is the least prime factor of a_n. Then $\sigma(a_n)/a_n$, $\phi(a_n)/a_n$ and $\psi(a_n)/a_n$ tend to* 1 *as $n\to\infty$.*

Proof. It is well known that for any integer $m>1$,

$$\frac{\sigma(m)}{m}=\prod_{p^\alpha\|m}(1-p^{-(\alpha+1)})\cdot\prod_{p|m}(1-p^{-1})^{-1}, \tag{3}$$

$$\frac{\phi(m)}{m}=\prod_{p|m}(1-p^{-1}) \quad \text{and} \quad \frac{\psi(m)}{m}=\prod_{p|m}(1+p^{-1}), \tag{4}$$

where the second product in (3) and the two products in (4) are taken over all prime factors p of m and the first product in (3) is taken over all p^α such that $p^\alpha|m$, $p^{\alpha+1}\nmid m$. Hence

$$1>\frac{\phi(a_n)}{a_n}=\prod_{p|a_n}(1-p^{-1})\geqq(1-q_n^{-1})^{\omega(a_n)}, \tag{5}$$

where $\omega(a_n)$ is the number of distinct prime factors of n. We have

$$a_n=\prod_{p^\alpha\|a_n}p^\alpha\geqq q_n^{\omega(a_n)}, \quad \text{so that}$$

$$\omega(a_n) \leqq \frac{\log a_n}{\log q_n}. \tag{6}$$

Hence from (5) and (6), $1 > \phi(a_n)/a_n \geqq (1 - q_n^{-1})^{\log a_n/\log q_n}$ so that

$$0 > \log\left(\frac{\phi(a_n)}{a_n}\right) \geqq \frac{\log a_n}{\log q_n}\log(1 - q_n^{-1}) > -\frac{2\log a_n}{q_n \log q_n}. \tag{7}$$

Since $\log a_n/(q_n \log q_n) \to 0$ as $n \to \infty$, it follows from (7) that $\log(\phi(a_n)/a_n) \to 0$, so that

$$\phi(a_n)/a_n \to 1 \quad \text{as} \quad n \to \infty. \tag{8}$$

We have from (3) and (4),

$$\frac{\sigma(a_n)}{a_n}\cdot\frac{\phi(a_n)}{a_n} = \prod_{p^\alpha \| a_n} (1 - p^{-(\alpha+1)}),$$

so that by (6),

$$1 > \frac{\sigma(a_n)}{a_n}\cdot\frac{\phi(a_n)}{a_n} > \prod_{p \mid a_n} (1 - p^{-2}) > (1 - q_n^{-2})^{\omega(a_n)}$$

$$> (1 - q_n^{-2})^{\log a_n/\log q_n}.$$

Hence

$$0 > \log\left\{\frac{\sigma(a_n)}{a_n}\cdot\frac{\phi(a_n)}{a_n}\right\} > \frac{\log a_n}{\log q_n}\log(1 - q_n^{-2}) > -\frac{2\log a_n}{q^2 \log q_n}. \tag{9}$$

Hence just as (8) followed from (7), it follows from (9) that

$$\frac{\sigma(a_n)}{a_n}\cdot\frac{\phi(a_n)}{a_n} \to 1 \quad \text{as} \quad n \to \infty. \tag{10}$$

Now, from (8) and (10), we obtain

$$\sigma(a_n)/a_n \to 1 \quad \text{as} \quad n \to \infty.$$

Since by (4), $(\phi(a_n)/a_n)\cdot(\psi(a_n)/a_n) = \Pi_{p \mid a_n}(1 - p^{-2})$, it follows by proceeding on the same lines as above that

$$\frac{\phi(a_n)}{a_n}\cdot\frac{\psi(a_n)}{a_n} \to 1 \quad \text{as} \quad n \to \infty. \tag{11}$$

Now, from (8) and (11), we obtain $\psi(a_n)/a_n \to 1$ as $n \to \infty$. Thus the theorem is proved.

REMARK 1. It may be noted that the condition $\log a_n/q_n \log q_n \to 0$ as $n \to \infty$ is equivalent to $(1 - q_n^{-1})^{\log a_n/\log q_n} \to 1$ as $n \to \infty$.

REMARK 2. It is known (cf. **[5]**, p. 343 or cf. **[7]**, p. 88) that every prime factor of F_n is of the form $2^{n+1}k + 1$, so that when $a_n = F_n$, we can take $q_n = 2^{n+1}k + 1$ and we see that

$$\frac{\log(2^{2^n} + 1)}{(2^{n+1}k + 1)\log(2^{n+1}k + 1)} < \frac{\log 2^{2^{n+1}}}{2^{n+1}\log 2^{n+1}} = \frac{2^{n+1}\log 2}{2^{n+1}(n+1)\log 2}$$

$$= \frac{1}{n+1} \to 0 \quad \text{as} \quad n \to \infty.$$

Also, it is known (cf. **[5]**, p. 247) that every prime factor of M_{p_n} is of the form $p_n k + 1$, so that when

$a_n = M_{p_n}$, we can take $q_n = p_nk + 1$ and we see that

$$\frac{\log(2^{p_n}-1)}{(p_nk+1)\log(p_nk+1)} < \frac{p_n \log 2}{p_n \log p_n} = \frac{\log 2}{\log p_n} \to 0 \quad \text{as} \quad n \to \infty.$$

REMARK 3. As examples of a_n other than F_n and M_{p_n}, we mention

(i) $a_n = a^{2^n} + 1$, where a is any even integer (cf. [5], p. 343)

(ii) $a_n = \frac{1}{3}(2^{p_n} + 1)$, $n > 2$ (cf. [5], p. 247) and more generally, $a_n = (a^{p_n} + 1)/(a + 1)$, $n \geqq 2$ and a is any integer > 1 (cf. [7], p. 87)

(iii) $a_n = (a^{p_n} - 1)/(a - 1)$, $n \geqq 2$ and a is any integer > 1 (cf. [7], p. 87 or cf. [6], p. 229) and

(iv) $a_n = p_n^{\alpha} \cdot b_n$, where $b_n = 1$ or b_n is an integer each of whose prime factors lie between p_n and $p_n{}^n$ with $\Omega(a_n) = \text{constant}$, $\Omega(a_n)$ being the total number of prime factors of a_n including multiplicity.

It may be noted that in example (i) above, every prime factor of a_n is of the form $2^{n+1}k + 1$ and in examples (ii) and (iii), every prime factor of a_n is either p_n or of the form $2p_nk + 1$.

REMARK 4. It can be easily seen that the converse of the above theorem is false. For example, if (i) $a_n = p_n^{\alpha p_n} (\alpha > 0)$, then $q_n = p_n$, $\log a_n / q_n \log q_n = \alpha \neq 0$ and (ii) $a_n = p_n^{np_n}$, then $q_n = p_n$, $\log a_n / q_n \log q_n = n \to \infty$; although, in both the cases $f(a_n)/a_n \to 1$ as $n \to \infty$, where $f = \sigma$, ϕ, or ψ.

In view of the theorem and Remark 4, the condition $\log a_n / q_n \log q_n$ on the sequence $\{a_n\}$ is only sufficient but not necessary in order that

$$\operatorname*{Lim}_{n \to \infty} \frac{f(a_n)}{a_n} = 1,$$

where $f = \sigma$, ϕ or ψ. So, it is natural to raise the following:

QUESTION 1. What are the necessary and sufficient conditions which the sequence $\{a_n\}$ must satisfy in order that $\operatorname{Lim}_{n\to\infty}(f(a_n)/a_n) = 1$, where $f = \sigma$, ϕ, or ψ?

Another question which might be worth raising in this connection is the following:

QUESTION 2. Suppose $\{a_n\}$ is a sequence satisfying the conditions of the theorem. Then what are all the arithmetical functions f for which $\operatorname{Lim}_{n\to\infty}(f(a_n)/a_n) = 1$?

References

1. Ranko Bojanić, Solution of Problem A4590, this MONTHLY, 62 (1955) 498–499.
2. E. Cohen, A property of Dedekind's ψ-function, Proc. Amer. Math. Soc., 12 (1961) 976.
3. L. E. Dickson, History of the Theory of Numbers, Vol. I (reprinted), Chelsea, New York, 1952.
4. P. Erdös, Problem A4590, this MONTHLY, 61 (1954) 350.
5. W. Sierpiński, Elementary Theory of Numbers, Polska Akademia Nauk, Monografie Matematyczne, Tom 42, Warszawa, 1964.
6. J. V. Uspensky and M. A. Heaslet, Elementary Number Theory, McGraw-Hill, New York, 1939.
7. I. M. Vinogradov, An introduction to the Theory of Numbers (translated from the sixth Russian edition), Pergamon Press, New York, 1955.

DEPARTMENT OF MATHEMATICAL SCIENCES, MEMPHIS STATE UNIVERSITY, MEMPHIS, TN 38152. (CURRENT ADDRESS: ANDHRA UNIVERSITY, WALTAIR, INDIA.)

MISCELLANEA

4. $2^{(5/2)^{(2/5)}}$ is a good approximation to e. See G. W. Brewster, Mathematical Gazette, 25 (1941) p. 49.

MATHEMATICAL EDUCATION

EDITED BY PAUL T. MIELKE AND SHIRLEY HILL

Material for this Department should be sent to Paul T. Mielke, Department of Mathematics, Wabash College, Crawfordsville, IN 47933, or to Shirley Hill, Department of Mathematics, University of Missouri, Kansas City, MO 64110.

CUPM ANNOUNCEMENT

The MAA's Committee on the Undergraduate Program in Mathematics (CUPM) is examining a possible recommendation that the standard collegiate mathematics major be revised to become a mathematical sciences major. The CUPM Panel considering this revision is seeking input from the mathematical community. The Panel is especially interested in hearing from mathematicians at schools that explicitly or implicitly now have a mathematical-sciences major (what do's and don't's can you recommend to other schools starting math sciences programs). Successful, innovative approaches to teaching various applied math courses are also sought. Please communicate your experiences and thoughts about this revision to Professor Alan Tucker, Department of Applied Mathematics and Statistics, State University of New York, Stony Brook, New York 11794, or Professor Richard Alo, Department of Mathematics, Lamar University, Beaumont, Texas 77710

A WAY MATHEMATICIANS CAN MEANINGFULLY CONTRIBUTE TO PRE-COLLEGE MATHEMATICAL TRAINING

JAMES H. JORDAN

Introduction. Prior to 1950, pre-college education was relatively free of direct influence of mathematicians, but indirect influences were strong. College prep was the main thrust of pre-college education, and admission to a college was judged on carefully documented student achievement.

During the 50's and 60's many reputable mathematicians began trying to directly influence pre-college mathematics education. Their motives were noble and based on the premise that curriculum changes in pre-college mathematics were long overdue. Names of many of these mathematicians can be found on various federally-financed writing teams, conference rosters, advisory committees, and project director's directories. The products and byproducts of these thrusts became known as the "new" math or the "modern" math. The success or non-success of these developments is widely debated.

A new avocation was found for many mathematicians, namely, retraining teachers to handle "new" math. This training was usually directed at the secondary teacher with a few isolated programs for the elementary teacher. Programs supported largely by N.S.F. were available in many colleges and universities in the country.

A "backlash" against many of these curriculums began to occur in the late 60's and early 70's. Certain indicators of mathematical competence began to descend. The general public and pre-college school systems began to look for "scapegoats" to blame for the descent. Prime candidates were these "carpetbaggers" called Mathematicians who had toyed with their curriculum. Mathematicians became about as popular with school systems as Golda Meir at a PLO convention. Many from outside mathematics seized upon this unpopularity to thrust their own theories into the curriculum.

Mathematicians were excluded from federal agencies and their committees that made decisions about development and direction in pre-college education. Many mathematicians curtailed their direct involvement and then retreated to the sidelines to spectate, or returned to the "dressing room" to pursue their initial research interests. At the Snowmass Conference of June 1973 it was observed "There was a concern expressed that schools are no longer looking toward mathematicians and mathematics educators at universities for assistance in their curriculum problems." **[1]**

Under the premise that many mathematicians could contribute meaningfully toward the solution of some of the problems of mathematics education at the pre-college level a proposal to the National Science Foundation was funded to conduct an experiment conditioning mathematicians for this noble purpose. The purpose of this paper is to report to the mathematical community the rationale for the project, the activities involved in the project, and the type of reactions we are getting about the project at the halfway point.

2. The conditioning program. As all mathematicians know, in order to solve a problem you must first find it, become intimately acquainted with it, understand the environment of the problem, the boundary conditions that prevail, and possess the necessary tools to attack it.

The place where these educational problems abound is in the classrooms of the pre-college schools and not in a university setting. Therefore, the conditioning process calls for the mathematicians to spend one full academic year in the pre-college classrooms with the teachers and students. For five days each week, six hours a day, the mathematician serves as a teacher's aide, helping the teacher tackle the everyday problem of teaching mathematics. Each level from kindergarten through senior high school is given equal emphasis. The teacher and the mathematician work as a team, each adding his or her own area of expertise to help the child learn mathematics. Since most mathematicians are woefully out of touch with the curriculum, the pedagogy, the theory of learning, and the operation of pre-college educational systems, it is vitally important to have a preparation period to acquaint them with what to expect when they begin their experience in the classrooms. This acquaintance period was designed to (i) load them with an extensive repertoire of appropriate activities to aid the students and impress the teachers and principals, (ii) study contrasting philosophies of education, learning theory, and educational psychology, (iii) examine educational problems imposed by the general public and the legislatures, (iv) accumulate knowledge of available and developing curriculums both "traditional" and "modern," and (v) coach them on methods of making teachers, principals, and school systems receptive to their assistance.

So in April of 1975 we had the mathematicians attend the fifty-third annual meeting of NCTM. Here they participated in numerous workshops, chatted with teachers, mathematics educational leaders, and each other, listened to speeches on current issues, and research reports concerning mathematics education, and observed the variety of curriculums and commercial products available. Then in July of 1975 they endured a four-week intensive program on the Washington State University campus. This included a course on illustrating mathematics concepts in the elementary school, a course on the structure and operation of large school systems, a graduate course on learning theory, and information about eight developed or developing mathematics education programs. (Those eight programs were (i) Secondary School Mathematics Curriculum Improvement Study — (SSMCIS); (ii) Comprehensive School Mathematics Program — (CSMP); (iii) Statistics by Example; (iv) Source Book on Applications; (v) The Arithmetic Project; (vi) The Madison Project; (vii) Unified Science and Mathematics in the Elementary School — (USMES); and (viii) Developing Mathematical Processes — (DMP).

In late August of 1975 the mathematicians were thrust into the classroom at various schools in the state and began their intensive exposure to pre-college education at the grass roots level. Those that survived the nine months in the classroom returned to Washington State University in June of 1976 for four additional weeks of courses. These courses were designed to give them techniques of assessment and evaluation, information about agencies that assist pre-college education and how to approach

them, awareness of additional developing mathematics programs, and a master's degree in education to add to their Ph.D. in mathematics.

At the conclusion of the program it was hoped that the mathematicians would be better equipped to assist pre-college education.

3. Interim report on the operation.

A. *Selecting the Participants.* There was some concern about the existence of a mathematician who would be willing to devote an entire year to this program. Matters were made even worse by an N.S.F. ruling that each participant would be limited to $100 a week living expenses with no salary.

In addition it was decided that we should not select participants from the ranks of the unemployed. These three negative restraints made the job of finding the participants quite difficult.

Utmost care was taken in the selection of participants. The selection procedure was begun in June of 1974 with wide advertisement of the program. Mathematics Departments around the country were contacted for possible faculty that were interested in the program. Announcements of the program were made in this MONTHLY, *Notices* of A.M.S. and in various newsletters of mathematics regional groups. Interested persons were invited to inquire about further information and application forms. About 125 persons made initial inquiries and thirty-seven completed application forms. When we received a completed application form and determined if the person was qualified to participate in the program, we sent letters to the chairperson, the divisional dean, the dean of the college of education, and the appropriate ranking administrator at the applicant's campus to make sure that the participant would not place tenure in jeopardy by participating in the program, to test the disposition of the school as to its intention of making adequate use of the newly-developed skills of the participant, and their intention of rewarding activities directed at pre-college education on a par with mathematical research and college teaching.

The author in most cases arranged a personal interview with each applicant to attempt to answer any further inquiries about the program, and to assess the applicant's potential to function at the pre-college level. Seventeen offers were made and the allocation of twelve participants accepted. Several others were informed that their participation could not be supported by project funds, but if they could arrange other financial support they could participate in the aspects of the program they desired. One person joined the project with his home institution covering all his expenses.

All thirteen of the participants hold positions in mathematics departments from around the country. The regional distribution is from Georgia (2), New York (2), Pennsylvania (2), Illinois, Missouri, New Mexico, Washington, West Virginia, and Wisconsin. They hold their Ph.D. degrees from Adelphi, SUNY-Stonybrook, Brown, UCLA, Oregon, Wyoming, Kentucky, Ohio State, LSU, Florida State, Georgia State, Pennsylvania State, and Washington State. Their mathematics research areas are Algebra (4), Number Theory (3), Analysis (2), Foundations (2), Differential Equations, and Applied Mathematics. The mean age is 33, and the mean number of years of college teaching experience is 6. Seven participants received considerable financial support from their home institution; one received partial support and five received no support.

We are very pleased with the participants in the program.

B. *Staffing.* A project of this kind is more likely to be successful if there is cooperation of several segments of the university community, the cooperation of the school districts, State Office of Education, and mathematics education organizations. Since mathematicians seldom listen to anyone other than mathematicians, we felt it was imperative that the project be directed by a mathematics professor, and administered through the mathematics department. In addition to the director, three others from the WSU Mathematics Department (C. T. Long, Jack M. Robertson, and D. W. DeTemple) were involved in teaching the courses. Two others from the WSU Mathematics Department (Don Bushaw and H. C. Wiser) presented special programs. One mathematics professor (W. E. Deskins) was imported to present a special program (CSMP). The education department was called upon to staff the courses in Advanced Learning Theory and in Education Evaluation and

Research Techniques; supervisors for the classroom experience portion oversee the Masters of Education degree, and assist in giving advice on extramural support for projects. The school systems contributed through their mathematics curriculum supervisors' coordination with the schools, resource persons for the courses on operating school districts and methods of teaching arithmetic, and some help in the presenting of special programs. The State Office of Education, through their mathematics supervisor, Elden Egbers, provided the assistance necessary to facilitate certification of the participants and coordinated several in-service programs for the mathematicians to practice their skills. The mathematics education organizations organized their regional programs to include workshops and talks by the participants. Two teachers from the Portland System were imported to present a special program that they had been using (USMES).

With all these factions helping the program, it was noticeably free of the kind of friction that sometimes occurs between two or more of the cooperating components. Indeed, there has been a significant improvement in the relations between the areas that have contributed to the program.

3. Reactions of the mathematicians. We had the mathematicians express their feelings about the valuable and worthless facets of the Denver NCTM meetings. Their most positive reaction was to each other. A sort of, *Esprit de Corps* evolved and the tone seemed to be that it was welcome to find others who were taking the same voyage. Mixed comments were given concerning aspects of the Denver program. Primary and intermediate workshops seemed to be the most favorably received, while the commercial exhibits seemed to lack relevance. Several commented favorably about meeting teachers and other mathematics education leaders; others responded negatively. Many expressed a desire to know more specifics about what their future experiences would be, and were unhappy that it was left ill-defined and vague. A general tone of "a worthwhile experience" was noticed, with a couple of dissenting opinions.

We also had them evaluate their summer training period. They were very unhappy about the red tape hangups at WSU and the inflexibility of the system to adapt to a special program of this nature. They seemed to have a preference for the mathematical methods course and the variety of curricula, and a dislike of the details of operating large school systems and the learning theory, although they acknowledged the usefulness of knowing the content of these courses.

Their reaction to the work in classrooms has been very positive at the primary and elementary level, a little less positive at the junior high level, and rather neutral at the high school level. Nearly all have expressed an enormous respect for elementary teachers that they had not had before.

They vividly relate their experiences in the schools and are beginning to bring up case histories and behavior of children, much the way teachers discuss children.

They feel they are now beginning to understand those extra forces that are imposed upon school systems that make changes slow and difficult. They are learning how to be more political with teachers and how to give constructive criticism without offending the person being criticized.

They enjoy giving workshops at the different meetings and special in-service days. They feel that they are now in a position to help with some aspects of improving mathematics education.

4. Reactions of the teachers. The high school teachers had widely differing reactions, from a negative "Who needs them?" to an enthusiastic "They provide an expertise for my students that I couldn't provide," or "It's wonderful to have some college professors really interested in finding what goes on here." When asked if they would accept another scholar to work in the classroom, they responded on an average between neutral and 'yes' without reservations. Of the 32 high school teachers polled, two were quite negative (emphatically 'no'), five were quite positive (would actively campaign for one). Some of the negative comments seemed to be sort of like a little league coach giving batting practice while Mickey Mantle was observing. Others seem to feel they were being evaluated by the administration. Still others felt the experience was superficial for the scholar and caused some dissension in the class. Those giving favorable comments felt the students benefited from

the experience, they were happy to have the extra help, they could use some ideas of the mathematical scholar, and it was professionally beneficial to have the mathematician with whom to converse. Most high school teachers believed that their opinion of mathematicians had improved, with only one claiming the experience had made him more antagonistic toward mathematicians.

The junior high school teachers were much more positive than the high school teachers. Of the 24 teachers polled, only one was quite negative and one was moderately negative, while seven were very positive. They responded a good strong, "Yes, we would do it again," without reservations. The negative responses came from a school that was very traditional and mimicked the conservative high school curriculum. The negative comments seemed to be, "Just another observer," "No way he could help me out," "Why didn't he teach more?" The positive responses were in the nature of helping with enrichment, having fresh ideas about remediation activities, helping in administering to the individual differences of students, helping with computer instruction and providing a resource person to help implement new ideas. Only one felt a deterioration of attitude toward mathematicians.

The elementary teachers were heavily in favor of the program. Of the 64 responding, only three had negative reactions and only one of these was quite negative, while 23 were extremely positive. In this area where classrooms are usually self-contained, rather than departmentalized, the teachers realized their weaknesses and were appreciative of any help. Their response to, "Would you have another mathematician help?" averaged midway from "Yes, without reservations" to "I would actively campaign for one." Those making negative comments felt that the inexperience of the mathematicians with younger children had seriously impaired their ability to operate effectively. Those making strong positive comments felt the variety of useful activities, the cooperative planning of lessons for individuals and small groups, the eagerness of the children to work with the mathematician and the vast mathematical knowledge of the individuals were extremely beneficial. Only one felt that their attitude toward mathematicians had deteriorated, while 27 felt that they had an improved attitude toward that species.

In general, the elementary teachers had had less exposure to mathematics, and were anxious to hear about improved techniques of instruction. The high school teachers were more mathematically informed and had a set method of delivery which did not allow for the flexibility of using the mathematician.

5. Reactions of the principals. Principals at all levels were much more positive than were the teachers. They, by a large majority, approved the program for the good of the scholars, the teachers (whether they want it or not), and the good of the students. No negative reactions seemed to ascend to the principal's office, with one exception. One school principal had noticed some friction between some of the teachers and the mathematician. This was the only principal responding who did not say they would actively campaign to have their school involved in a repeat of the program. Most principals had a better feeling about mathematicians after their experience with the program.

6. Reactions of the mathematics curriculum people. The mathematics curriculum people from the involved districts continue to laud the program. They express the feeling that good things are happening that are beneficial to all concerned.

7. Unexpected beneficial windfalls. Three classes on metric education strategies have been conducted by the mathematicians with excellent feedback from the teachers. One class on the use of manipulatives in elementary school mathematics has been conducted by the mathematicians. The Pacific Science Center has utilized the mathematicians in their mathematics classes for school children from the area. The University of Washington has used one of the mathematicians to handle a remedial arithmetic class for freshmen. The migrant education center for Washington has used one of our Spanish-speaking mathematicians to build mathematical games with instructions in Spanish to be printed and distributed through the migrant center. Workshops were held by two mathematicians for

the Native American teachers aides of the Yakima tribe. Workshops for kindergarten teachers on mathematics manipulatives appropriate for the very young were held in the Spokane school district. The Seattle school districts developed four impact area schools that inservice and supply teachers with necessary materials and software for activities. The mathematicians worked the impact schools and provided the inservice for the teachers. Seattle also developed a central drop-in inservice and activity center for counselling teachers and developing the activities for next day lessons. Two mathematicians manned these centers after teaching hours. At a state inservice day the mathematicians provided 50% of the Puget Sound Mathematics Council's Program for the teachers of the area. The Lawrence School of Science in cooperation with Texas Instrument has invited all of the mathematicians in the program to an awareness workshop on their newly-developed curriculum that utilizes pocket calculators.

8. Tentative conclusion. During a time when there seem to be more mathematicians around to do what mathematicians have done in the past, it seems reasonable to investigate areas where mathematicians can meaningfully contribute. If their contribution is sought by society, then by providing this service they will improve the reputation of the profession. We believe we are showing that pre-college education is an excellent area for mathematicians to dispense their expertise, provided they undergo some retraining. We hope we are showing one kind of retraining that brings the desired results. Persons interested in more specifics concerning the training should contact the author.

The author was partially supported by NSF Grant GW–7919.

Reference

1. Report on the Conference on the K-12 Mathematics Curriculum, Mathematics Educational Development Center of Indiana University, (1973), p. 6.

DEPARTMENT OF MATHEMATICS, WASHINGTON SATE UNIVERSITY, PULLMAN, WA 99164.

MATHEMATICS FOR SOCIAL SCIENCE STUDENTS

J. S. HARTZLER

1. Introduction. Mathematics has for centuries been a major component in the education of scientists. In the last several decades the importance of mathematics in the education of business majors has expanded dramatically. Many colleges have been offering a mathematics course sequence for students of business and social science. This practice was consistent with the availability of textbooks allegedly written for students of business and social science. Because of the increasing emphasis on mathematics in the curriculum for business students, these dual purpose courses have gravitated toward meeting the needs of the business students by relating the mathematics taught primarily to economics and consequently losing the interest and attention of social science students interested in psychology, anthropology, political science and sociology. Social science students have often moved, with the encouragement of their advisors, to "liberal arts" mathematics courses, which are usually general enough to accommodate the heterogeneous clientele they attract and hence do little to relate mathematics to the particular interests of the social science students. The solution to the problem may be in a course sequence designed especially for social science students.

2. Course content. A course sequence for social science students should have three major objectives:

1. to develop competence with basic mathematical tools,

2. to demonstrate the utility of mathematics in the social sciences, and
3. to prepare students to understand mathematical models in the current social science literature.

T. L. Saaty, in his book (see [3]) suggests that set theory, graph theory, matrix theory, statistics, stochastic processes, and computer simulation, are the topics most likely to find utility in the social sciences and hence, an introduction to these topics would be appropriate in the training of social scientists. Additionally, some consideration of polynomial and exponential functions would be appropriate. The absence of calculus is a significant argument for the separation of the mathematics preparation for business and social science students, since analysis is used heavily in economics.

3. A modelling approach. If a course sequence for social scientists is to meet the objectives listed above, the approach to teaching the course must be quite different than has been the case in many so called "service" mathematics courses in the past. In particular, the mathematics instructor must be willing to get involved in relating mathematics to social science so that he or she can assist students in understanding how mathematical models are used to solve social science problems. Students need to learn much more than mathematical solutions to mathematical problems. They need to struggle with the difficult task of developing, interpreting, and validating mathematical models for "real" problems. A modelling approach to mathematics will be a major asset in overcoming the lack of interest in or fear of mathematics often displayed by social science students.

There are many interesting applications of set theory, graph theory, and matrix theory that can be developed in a first course for social science students with high school algebra as the sole prerequisite. For example, counting techniques in survey analysis, Harary measure of status in a hierarchy, break-even analysis, demographic projections, dominance in social organizations, communications analysis, cliques in a social structure, measuring social mobility, predicting voter behavior, the Hare system of counting ballots, and analyzing the spread of a rumor.

In addition to a first course including the topics mentioned above and a course in applied probability and statistics, a third course to consider mathematical models in the current social science literature would be appropriate for students who expect to pursue an advanced degree. Such a course should involve both mathematicians and social scientists in the instructional process. Additionally, this course would provide an interesting elective for students of the mathematical sciences who are interested in studying the utility of mathematics in areas other than the physical sciences and economics. If both types of students enroll in the course, it would be appropriate to involve several students from each discipline in research cells for the purpose of jointly attempting to develop a mathematical model for a "real" problem.

4. Implementation. An effective educational program in mathematics for social science students can be developed and implemented if and only if several mathematicians are willing to familiarize themselves with mathematical models in the current social science literature. On the surface it may appear that the reward for this effort would only be the opportunity to teach a course in which the mathematics is not too sophisticated. However, there is great potential for mathematics to play an increasingly important role in the social sciences in future decades and therein lies the possibility of professional rewards for the adventuresome mathematician.

References

1. R. W. Brewster, Government in Modern Society, Houghton–Mifflin, Boston, 1963.
2. A. Mizrahi and M. Sullivan, Finite Mathematics with Applications, Wiley, New York, 1973.
3. T. L. Saaty, Topics in Behavioral Mathematics, MAA, Washington, D. C., 1973.

DEPARTMENT OF MATHEMATICS, PENNSYLVANIA STATE UNIVERSITY, CAPITOL CAMPUS, MIDDLETOWN, PA 17057.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before February 28, 1978.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

E 2677. *Proposed by Erwin Just, Bronx Community College, C.U.N.Y.*

Let $n \geqq 2$ be an integer. Show that there exists a function $f: \mathbf{R} \to \mathbf{R}$ such that $f(x)+f(2x)+\cdots+f(nx)=0$ for all x and $f(x)=0 \Leftrightarrow x=0$.

E 2678. Proposed by Edward T. H. Wang, Wilfrid Laurier University, Waterloo, Ontario, Canada*

Find the maximum number of ones in an $n \times n$ $(0,1)$-matrix whose square is again a $(0,1)$-matrix.

E 2679. *Proposed by Solomon W. Golomb, University of Southern California*

If a positive integer m has a prime factor greater than 3, show that $4^m - 2^m + 1$ is composite.

E 2680*. *Proposed by Jerrold W. Grossman, Oakland University, Michigan*

Let $ABCD$ be a convex quadrilateral in the hyperbolic plane. Assume that $AD = BC$ and that

$$\measuredangle A + \measuredangle B = \measuredangle C + \measuredangle D.$$

Does $AB = CD$ follow from the above hypotheses? (It does in the Euclidean plane.)

E 2681. *Proposed by David Burman, Bell Laboratories, Holmdel, New Jersey*

If $x + y = 1$ show that

$$\sum_{i=0}^{m-1} \binom{n+i-1}{i} x^i y^n + \sum_{j=0}^{n-1} \binom{m+j-1}{j} x^m y^j = 1.$$

E 2682. *Proposed by Douglas Hensley, University of Illinois at Urbana-Champaign*

Let E be an ellipse in the plane whose interior area $A \geqq 1$. Prove that the number n of integer points on E satisfies $n < 6A^{1/3}$.

SOLUTIONS OF ELEMENTARY PROBLEMS

An Iterated Function

E 984 [1951, 564; 1952, 252; 1976, 567]. *Proposed by Joseph Rosenbaum.*

(a) Find $f(x)$ when $f[f(x)]=x^2-2$.

(b) More generally, find $f_1(x)$ when $f_n(x)=x^2-2$, where $f_n(x)$ is defined by the relation $f_{r+1}(x)=f_1[f_r(x)]$.

II. *Comment by Bruce Reznick, Duke University.* We show that (a) has no solutions if we insist that f is defined on the whole real line.

Let $g(x)=x^2-2$ and suppose that $f_2(x)=g(x)$. Put

$$h(x)=g_2(x)=g[g(x)]=x^4-4x^2+2.$$

The fix-points of g are -1 and 2. The set of fix-points of h is $S=\{-1,2,(-1\pm\sqrt{5})/2\}$. It is clear that $a,b\in S$ and $f(a)=f(b)$ imply that $a=b$. Indeed, $f(a)=f(b)$ implies $f_4(a)=f_4(b)$, i.e., $h(a)=h(b)$, and we also have $h(a)=a$ and $h(b)=b$ since $a,b\in S$. If $a=-1$ or 2 then $g[f(a)]=f_3(a)=f[g(a)]=f(a)$ and consequently $\{f(-1),f(2)\}=\{-1,2\}$.

For $a=(-1+\sqrt{5})/2$ we have $h[f(a)]=f_5(a)=f[h(a)]=f(a)$. Thus $f(a)\in S$. Since f induces a bijection $S\to S$ and $g(a)=a^2-2\neq a$ implies $f(a)\neq a$, we must have $f(a)=b=(-1-\sqrt{5})/2$. It follows that $f(b)=a$ and we have a contradiction $a=f(b)=f[f(a)]=g(a)$.

This negative result for (a) implies that (b) has no solution on the whole real line if n is even. The case of odd $n\geqq 3$ remains open (but see comments below for $n=3$).

The above negative result persists if the equation (a) is considered in a set D (of real or complex numbers) containing S.

Comments. Rufus Isaacs observes that g has only one 2-cycle $\{(-1\pm\sqrt{5})/2\}$ and then deduces the above results from his paper *Iterates of fractional order*, Canad. J. Math. 2(1950), 409-416. He also discusses several cases when f is defined on a subset D of the complex numbers. If D contains the interval $(-2,2]$ he shows that part (b) has no solution for $n=3$.

Counting Points in a Configuration

E 2593 [1976, 378]. *Proposed by Jeanne W. Kerr and John E. Wetzel, University of Illinois.*

Three points are given on each of three parallel lines, the three lines not all lying in the same plane. These points by threes, one on each line, determine 27 triangular plates and these triangular plates by threes could, on the face of it, meet to determine as many as $\binom{27}{3}=2925$ points, though it is clear that not that many can actually occur. At most how many points can the 27 plates determine?

Solution by the proposers (revised by the editor). Let X, Y, Z be the given parallel lines and Π the infinite prism having these lines as edges. A *plate* is a closed triangle having one vertex on each of the lines X, Y, Z. A *configuration* Σ is a finite set of plates which has the following property: If T is a plate and every vertex of T is a vertex in some plate in Σ then $T\in\Sigma$. From now on Σ will denote a configuration. A *vertex of* Σ is a vertex of a plate in Σ. Let p, q, r be the numbers of distinct vertices of Σ on X, Y, Z (in some order), with $p\leqq q\leqq r$. Then we say that Σ is of *type* (p,q,r).

Let $\pi(\Sigma)$ be the set of points P inside Π such that there exist three plates in Σ whose intersection is $\{P\}$. Clearly, $\pi(\Sigma)$ is finite and we put $N(\Sigma)=|\pi(\Sigma)|$.

Our problem is to find the maximum of $N(\Sigma)$ when Σ varies over all configurations of type $(3,3,3)$. If Σ is an optimal configuration of that type then it is clear that no 4 plates of Σ can meet at a point inside Π, and no 3 plates of Σ can meet at a point in the interior of a face of Π. From now on we assume that Σ is a configuration of type $(3,3,3)$ and that these degeneracies do not occur.

Let Ω be a subconfiguration of Σ. For $P\in\pi(\Omega)$ let Ω_P be the smallest configuration containing all

the plates of Ω which contain P. Let $\pi_0(\Omega)$ be the subset of $\pi(\Omega)$ consisting of those points P for which $\Omega_p = \Omega$. It is clear that

$$N(\Sigma) = \Sigma N_0(\Omega), \qquad N_0(\Omega) = |\,\pi_0(\Omega)\,|,$$

where the sum is over all subconfigurations $\Omega \subset \Sigma$. If Ω is of type (2,2,2), (2,2,3) or (2,3,3), then $N_0(\Omega)$ just depends on the type and we have

$$N_0(2,2,2) = 2, \quad N_0(2,2,3) = 6, \quad N_0(2,3,3) = 9.$$

This can be verified by inspection. For instance, to verify $N_0(2,3,3) = 9$, we label the points on X, Y, Z by $A_1, A_2; B_1, B_2, B_3; C_1, C_2, C_3$, respectively, listed from the left to the right. A plate $A_iB_jC_k$ will be abbreviated by ijk. Then the 9 points in $\pi_0(\Omega)$ are the ones determined by the following triples of plates:

113, 122, 231;	113, 132, 221;	122, 131, 213;
122, 213, 231;	123, 131, 212;	123, 132, 211;
123, 212, 231;	132, 213, 221;	133, 212, 221.

Since Σ contains 27 configurations of type (2,2,2), 27 configurations of type (2,2,3), and 9 configurations of type (2,3,3), we find that

$$N(\Sigma) - N_0(\Sigma) = 27 \cdot 2 + 27 \cdot 6 + 9 \cdot 9 = 297.$$

Let A_i, B_i, C_i $(i = 1,2,3)$ be the vertices of Σ with A_i on X, B_i on Y, and C_i on Z, listed in the same order. Let P, Q, R, be the points

$$B_1C_3 \cap B_3C_1, \qquad C_1A_3 \cap C_3A_1, \qquad A_1B_3 \cap A_3B_1,$$

respectively. One can check that if the triangle PQR meets the plate $A_2B_2C_2$ then $N_0(\Sigma) = 10$ and otherwise $N_0(\Sigma) = 8$. In the case when $N_0(\Sigma) = 10$, the points in $\pi_0(\Sigma)$ are determined by the following triples of plates:

113, 222, 331;	113, 232, 321;	122, 213, 331;
122, 231, 313;	123, 231, 312;	123, 232, 311;
131, 222, 313;	131, 223, 312;	132, 213, 321;
132, 223, 311.		

Thus $\max N(\Sigma) = 297 + 10 = 307$.

A number of incorrect solutions were submitted.

For more details see the paper by the proposers in *Geometriae Dedicata* 4(1975), 279–289.

Erratic Behavior of the Totient Function

E 2599 [1976, 482]. *Proposed by Bernardo Recamán S., Bogotá, Colombia*

Are there arbitrarily large positive integers N such that for all $n \geqq N$ we have $\varphi(n) \geqq \varphi(N)$ while $\varphi(n) \leqq \varphi(N)$ when $n \leqq N$? (φ denotes Euler's totient function.)

Solution by Lael F. Kinch, University of Louisville. Let $N \geqq 5$. We make use of Bertrand's postulate: For any integer $m \geqq 4$, there is a prime strictly between m and $2m - 2$.

Hence, there is a prime p such that

$$\left[\frac{N+3}{2}\right] < p < 2\left[\frac{N+3}{2}\right] - 2 \leqq N + 1.$$

We have

$$\varphi(p) = p - 1 > \frac{N+1}{2} \geqq \varphi(N+1) \qquad (N \text{ odd}), \qquad\qquad \varphi(p) = p - 1 > \frac{N}{2} \geqq \varphi(N) \qquad (N \text{ even}).$$

Thus N does not have the required property.

Since $\varphi(1)=\varphi(2)=1$, $\varphi(3)=\varphi(4)=2$ and $\varphi(n)\geqq 2$ for $n\geqq 3$. The integers 1, 2, 3, 4 have the required property.

Also solved by E. J. Barbeau (Canada), D. M. Bloom, Robert Breusch, Brian Conrey, Lorraine Foster, Marguerite Gerstell, Gustaf Gripenberg (Finland), Anita Grossman, I. M. Isaacs, Eli Isaacson, Wells Johnson, Elgin Johnston, Mark Kleiman, Margret Kothman, Vojtech Lásló (Czechoslovakia), Jordon Levy, L. E. Mattics, M. Ram Murty & V. Kumar Murty (Canada), Charles Nicol, R. W. K. Odoni (England), Barry Powell, Chester Palmer, V. Sita Ramaiah (India), Bruce Richmond (Canada), Eric Rosenthal, Robert Shafer, Nathaniel Sharpe, and Lou Thurston.

Minimum Modulus for a Polynomial

E 2600 [1976, 482]. *Proposed by Ron Evans, University of Wisconsin*

Fix $r\geqq 2$ and suppose that z_1, z_2, z_3, z_4 are complex numbers of modulus $\geqq r$. Find the point at which

$$2-(z_1+z_2)(z_3+z_4)+z_1z_2z_3z_4$$

attains its minimum modulus.

Solution by Elgin Johnston, student, University of Illinois. The minimum modulus is attained when $z_1=z_2=\bar{z}_3=\bar{z}_4$, and $|z_1|=r$. Replacing z_i by $1/w_i$, it suffices to show that $[2w_1w_2w_3w_4-(w_1+w_2)(w_3+w_4)+1]/w_1w_2w_3w_4$ attains its minimum modulus for $|w_i|\leqq 1/r$ when $w_1=w_2=\bar{w}_3=\bar{w}_4$, $|w_1|=1/r$. This is clear for the factor $(w_1w_2w_3w_4)^{-1}$. As for the polynomial in parentheses,

$$\begin{aligned}&|2w_1w_2w_3w_4-(w_1+w_2)(w_3+w_4)+1|\\&=|(1-w_1w_3-w_1w_4)(1-w_2w_4-w_2w_3)-w_1w_2(w_3^2+w_4^2)|\\&\geqq\left(1-\frac{2}{r^2}\right)^2-\frac{2}{r^4}.\end{aligned}$$

This lower bound, which is positive if $r\geqq 2$, is in such a case attained precisely when $w_1=w_2=\bar{w}_3=\bar{w}_4$, $|w_1|=r$.

Also solved by E. J. Barbeau (Canada), L. E. Mattics, L. A. Shepp, and the proposer.

Binomial Sum and Legendre Polynomials

2601 [1976, 482]. *Proposed by Robert Weinstock, Oberlin College*

Prove that

$$\sum_{k=0}^{[n/2]}\frac{\binom{2n-k}{k}}{\binom{2n-k}{n}}\,\frac{2n-4k+1}{2n-2k+1}\,2^{n-2k}=1.$$

I. *Solution by Otto G. Ruehr, Michigan Technological University.* We shall prove a more general identity

$$(1)\qquad \sum_{k=0}^{m}\frac{\binom{2n-k}{k}}{\binom{2n-k}{n}}\,\frac{2n-4k+1}{2n-2k+1}\,2^{n-2k}=\frac{\binom{n}{m}2^{n-2m}}{\binom{2n-2m}{n-m}}.$$

If $m = [n/2]$, i.e., $n = 2m$, or $n = 2m+1$, then the right hand side of (1) equals 1.

To prove (1), let $a(n,k)$ be the general term under the summation sign in (1) and let $b(n,m)$ be the right hand side of (1).

Obviously, (1) is true for $m = 0$. We prove that it is true for all m such that $2m \leqq n$ by using induction on m and the relation

$$a(n, m+1) = b(n, m+1) - b(n, m)$$

which is easy to verify.

II. *Solution by Elizabeth A. McHarg, University of Glasgow and Richard A. Groeneveld, Iowa State University (independently).* One has

$$x^n = \sum_{k=0}^{[n/2]} c_{2k} P_{n-2k}(x),$$

where

$$c_{2k} = \frac{\binom{2n-k}{k}}{\binom{2n-k}{n}} \frac{2n-4k+1}{2n-2k+1} 2^{n-2k}$$

and $P_k(x)$ is the Legendre polynomial of degree k.

The required identity is obtained by putting $x = 1$, since $P_k(1) = 1$ for all k. For the above expansion of x^n in terms of the Legendre polynomials see, e.g. G. Sansone, *Orthogonal Functions*, p. 194 (Interscience 1959).

Also solved by Kendall Barker, J. C. Binz (Switzerland), M. T. Bird, L. Carlitz, Sylvan Greene, Eli Isaacson, L. E. Mattics, Ram Murty & Kumar Murty (Canada), Robert Shafer, Franklin Smith, David Wright, and the proposer.

Editor's Comment. There was an error in the printed version of this problem [1976, 482], namely $2n - 4k + 1$ was printed as $2n - 4k - 1$. All solvers (with one exception) corrected this obvious misprint.

A Combinatorial Identity

E 2602 [1976, 482]. *Proposed by C. L. Mallows, Bell Laboratories, Murray Hill, New Jersey*

Prove that

$$\sum_{i=0}^{a-1} \binom{b+i-1}{b-1} \binom{2n-b-i}{n-b} = \sum_{i=b}^{n} \binom{a+i-1}{a-1} \binom{2n-a-i}{n-a}.$$

Solution and generalization by John W. Pratt, Harvard University. More generally, we have

$$\sum_{i=0}^{a-1} \binom{b+i-1}{b-1} \binom{m-b-i}{n-b} = \sum_{i=b}^{m-n} \binom{a+i-1}{a-1} \binom{m-a-i}{m-n-a}, \tag{1}$$

which contains the given identity as a special case ($m = 2n$).

The left hand side counts integral sequences $(k_1, k_2, \ldots, k_n)$ such that $1 \leqq k_1 < k_2 < \cdots < k_n \leqq m$ and $k_b \leqq a + b - 1$. The right hand side counts integral sequences $(s_1, s_2, \ldots, s_{m-n})$ such that $1 \leqq s_1 < s_2 < \cdots < s_{m-n} \leqq m$ and $s_a \geqq a + b$.

Each sequence $(k_1, \ldots, k_n)$ determines a unique complementary sequence $(s_1, \ldots, s_{m-n})$ such that $k_1 < \cdots < k_n$, $s_1 < \cdots < s_{m-n}$ and

$$\{k_1, \ldots, k_n, s_1, \ldots, s_{m-n}\} = \{1, 2, \ldots, m\}.$$

In that case we have

$$k_b \leqq a + b - 1 \Leftrightarrow s_a \geqq a + b.$$

This implies that the two members of (1) are equal.

Also solved by Einar Andresen (Norway), J. C. Binz (Switzerland), M. T. Bird, D. M. Bloom, L. Carlitz, Sylvan Greene, William Guenther, Eli Isaacson, Elgin Johnston, Ram Murty & Kumar Murty (Canada), Kelev Pedro (Israel), E. J. F. Primrose (England), Reinhart Razen (Austria), S. M. Samuels, Franklin Smith, Paul Vojta, David Wright, and the proposer.

A Symmetric Sum Inequality

E 2603 [1976, 483]. *Proposed by M. S. Klamkin, University of Alberta*

Let $x_i > 0$ $(1 \leqq i \leqq n)$. Prove that

$$r \cdot \sum \frac{x_1 x_2 \cdots x_r}{x_1 + x_2 + \cdots + x_r} \leqq \binom{n}{r} \left(\frac{x_1 + \cdots + x_n}{n} \right)^{r-1}$$

and that equality holds if and only if $x_1 = x_2 = \cdots = x_n$. (The "symmetric sum" above consists of $\binom{n}{r}$ terms.)

Solution by Lawrence A. Shepp, Bell Laboratories, Murray Hill, New Jersey. If $x_i \neq x_j$ then replacing x_i and x_j by $\frac{1}{2}(x_i + x_j)$ increases the left hand side and leaves the right hand side constant as is easy to see. Thus the maximum of the left side under fixed $x_1 + \cdots + x_n$ occurs only for $x_1 = \cdots = x_n$ in which case a direct calculation shows that the equality holds.

Also solved by E. J. Barbeau (Canada), Peter de Buda, Brian Conrey, Thomas Foregger, Siegfried Gabler (Germany), Zoárd Geöcze (Brazil), M. G. Greening (Australia), Richard Groeneveld, Ellen Hertz, Mark Kleiman, Joel Levy, Henry Lieberman, S. C. Locke (Canada), Russell Lyons, L. E. Mattics, Ram Murty & Kumar Murty (Canada), R. J. Serfling, Lou Thurston, John Tung, Paul Vojta and the proposer.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Hill Center, Rutgers University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate signed sheets and should be mailed before February 28, 1978.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

6174. *Proposed by Bartell W. Huff, Queen's University, Kingston, Ontario*

A family $\mathscr{F} = \{X_\lambda \mid \lambda \in \Lambda\}$ of random variables is said to be uniformly integrable if

$$\lim_{\alpha \to \infty} \sup_{\lambda} E\,|X_\lambda| \cdot I_{[|X_\lambda| \geqq \alpha]} = 0,$$

where I_A is the indicator function of the event A. One sufficient condition for uniform integrability is that there exists a random variable Y such that $|X_\lambda| \leqq Y$ a.s., $\forall \lambda$, and $EY < \infty$. Problem 6085 [1976, 292] asks whether the converse is true.

A weaker sufficient condition is that there exists a nonnegative random variable Y such that $P[|X_\lambda| \geqq \alpha] \leqq P[Y \geqq \alpha]$, $\forall \alpha > 0$, $\forall \lambda$ and $EY < \infty$ (Billingsley, *Convergence of Probability Measures*, p. 32). Is the converse to the weaker condition true?

6175. *Proposed by Ignacy I. Kotlarski, Oklahoma State University*

Let $(X_1, X_2, \ldots, X_n)$ be an n-dimensional real random vector. Consider the random polynomial of order n $(n = 2, 3, \ldots)$ on the complex plane,

$$P_n(\lambda) = (\lambda - X_1)(\lambda - X_2) \cdots (\lambda - X_n), \quad \lambda \in Z$$

and define

$$Z_n = \frac{1}{i}\frac{P_n(i)-(-1)^n P_n(-i)}{P_n(i)+(-1)^n P_n(-i)}, \qquad n=2,3,\ldots.$$

Show that if one of the X_k is independent from the others and follows the Cauchy distribution

$$P(X_k \leqq x) = \frac{1}{2} + \frac{1}{\pi}\arctan x, \qquad x \in R,$$

then all the Z_n are real random variables having the same Cauchy distribution.

6176. *Proposed by Morris Newman and Daniel Shanks, National Bureau of Standards.*

Prove that for the most common type of simple group, which is designated $\mathrm{PSL}_2(p^n)$, its order N is never a perfect square. Find at least one simple group that does have square order.

6177. *Proposed by Adrian R. Wadsworth, University of California, San Diego, La Jolla.*

Let K be a perfect field of prime characteristic. Prove that if R is a Noetherian integral domain with quotient field K then $R = K$.

6178*. *Proposed by Robert Kowalski, Winona, Minnesota.*

Define the shape of a rectangle to be the ratio of the longer side to the shorter side. Suppose one has an unlimited number of congruent squares at one's disposal. Given shape α and an error ε, what is the least number of squares one needs to construct a rectangle whose shape differs from α by less than ε?

6179. *Proposed by E. Ehrhart, University of Strasbourg, France.*

Find all cubes in a cubic lattice whose vertices are lattice points.

SOLUTIONS OF ADVANCED PROBLEMS

Positive Definite Hermitian Matrix

6072 [1976, 140]. *Proposed by Wayne Lawton, Institute for Advanced Study*

Let $a_1,\ldots,a_n$ be n distinct complex numbers such that $0<|a_k|<1$ for $1\leqq k\leqq n$. Let $B=(b_{ij})$ be the $n\times n$ Hermitian matrix defined by $b_{ij}=a_i\bar{a}_j/(1-a_i\bar{a}_j)$ for $1\leqq i,j\leqq n$. Then B is positive definite and the following equality is valid

$$\max_{x_i\in\mathbb{C}}\left\{|x_1+\cdots+x_n|^2\colon \sum_{1\leqq i,j\leqq n} b_{ij}x_i\bar{x}_j = 1\right\} = \prod_{k=1}^{n}|a_k|^{-2}-1.$$

Solution by Finbarr Holland, University College, Cork, Ireland. Let

$$\phi(z)=\prod_{k=1}^{n}\frac{z-a_k}{1-\bar{a}_k z} \qquad (|z|\leqq 1)$$

and $K=\phi H^2$, where H^2 stands for the Hardy space of square-integrable analytic functions on the unit circle with inner-product

$$\langle f,g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(e^{it})\overline{g(e^{it})}\,dt.$$

Consider the linear functional $f\to f(0)$ on $K^\perp$, the orthogonal complement of K. Since

$$f(0)=\langle f,1\rangle=\langle f,1-g\rangle, \qquad (f\in K^\perp,\ g\in K),$$

the norm of this functional is given by $\min\{\|1-g\|: g \in K\} = (1-|\phi(0)|^2)^{\frac{1}{2}}$, the latter being the norm of $1-\overline{\phi(0)}\phi$, the orthogonal projection of 1 onto $K^\perp$. Thus

$$(*) \qquad |f(0)|^2 \leqq \|f\|^2(1-|\phi(0)|^2) \qquad (f \in K^\perp)$$

and equality holds if and only if $f = \lambda\ (1-\overline{\phi(0)}\phi)$, where λ is constant.

To relate these remarks to the problem, note that any $f \in K^\perp$ is of the form

$$f(z) = \sum_{k=1}^{n} \frac{x_k}{1-\bar{a}_k z} \qquad (|z| \leqq 1),$$

where the x_k are constants. For such f, $f(0) = \sum_{k=1}^{n} x_k$ and

$$\|f\|^2 = \sum_{s=0}^{\infty} \left| \sum_{k=1}^{n} x_k \bar{a}_k^{\,s} \right|^2 = \left| \sum_{k=1}^{n} x_k \right|^2 + \sum_{1 \leqq i,j \leqq n} \bar{x}_i x_j b_{ij}.$$

Thus (b_{ij}) is positive definite.

Using now the inequality $(*)$ it follows that

$$\left| \sum_{k=1}^{n} x_k \right|^2 \leqq \left(\sum_{1 \leqq i,j \leqq n} \bar{x}_i x_j b_{ij} \right) (|\phi(0)|^{-2} - 1),$$

and that equality occurs if the x_k are chosen so that

$$\lambda(1-\overline{\phi(0)}\,\phi(z)) = \sum_{k=1}^{n} \frac{x_k}{1-\bar{a}_k z} \qquad (|z| \leqq 1).$$

The stated result now follows.

Also solved by C. F. Schubert (Canada).

Singular Monotonic Functions

6073 [1976, 140]. *Proposed by George Crofts, Virginia Polytechnic Institute and State U*

Let f be an increasing real-valued function from $[a, b]$ onto $[c, d]$ and let m denote Lebesgue measure. If there is a set $E \subseteq [a, b]$, with $m(E) = 0$, for which $m(f(E)) = d - c$, must f be singular (i.e., $f' = 0$ almost everywhere)?

Solution by L. E. Mattics, University of South Alabama. It is known, of course, that f is continuous and differentiable a.e. For convenience assume that $a, b \in E$. If $f'(x) \geqq M > 0$ on a set with measure $r > 0$, then $f'(x) \geqq M$ on a closed set $B \subseteq [a, b] - E$ with $m(B) \geqq r/2$. If $x \in B$ then there is a $\delta_x > 0$ such that $(f(w) - f(x))/(w - x) \geqq M/2$ if $|w - x| < \delta_x$.

Let $\varepsilon > 0$ be given, $f(E)$ and $f(B)$ are disjoint; there is a family $I(t)$ of open intervals such that $\Sigma l(I(t)) < \varepsilon$ and $f(B) \subseteq \cup I(t)$. If $x \in B$, there is a t_x such that $f(x) \in I(t_x)$. Since $f'(x) \geqq M$ (and f is continuous at x), we can pick a_x and b_x in $[a, b]$ such that

$$f(x) \in (f(a_x), f(b_x)) \subseteq I(t_x)$$

and $a_x < x < b_x$ with $x - a_x < \delta_x$ and $b_x - x < \delta_x$. Then $f(b_x) - f(a_x) \geqq M(b_x - a_x)/2$. For each x in B we pick such an interval (a_x, b_x) and, because B is closed, we take a minimal finite subcovering $J_1, \ldots, J_n$. Since this is a minimal subcovering, no point in B is in more than two of these intervals, so

$$2\varepsilon > 2\Sigma l(I(t)) \geqq M\Sigma l(J_i) \geqq Mm(B) \geqq Mr/2.$$

This proves that $f'(x) = 0$ almost everywhere on $[a, b]$.

Also solved by D. R. Arterburn & J. L. Winter, Peter Borwein (Canada), R. A. Christiansen, Roy Davies

(England), Max Garbutt, M. B. Gregory, A. A. Jagers (Netherlands), Nicholas Passell, and the proposer.

Notes. The solutions of Jagers and Arterburn–Winter reduce the problem to an exercise (17.26) in Hewitt and Stromberg, *Real and Abstract Analysis*. Davies and Christiansen reduce the problem to Problem 6074 below, showing that f has arc length 2 (once (a,b), (c,d) have been taken as $(0,1)$). The proposer remarks that the problem is somewhat parallel to the known theorem: If f is continuous, real-valued, of bounded variation, and if the image of each null set is a null set, then f is absolutely continuous.

Length of Arc of a Monotonic Function

6074 [1976, 140]. *Proposed by H. L. Montgomery, University of Michigan*

Let f be a weakly increasing continuous function defined on $[0,1]$ with $f(0)=0$, $f(1)=1$ and let L denote the arc length of the curve $(x,f(x))$, $0 \leqq x \leqq 1$. Prove that $L \leqq 2$, with equality if and only if $f'(x)=0$ almost everywhere. (Compare Problem 6007 [1976, 663] which gives a partial result.)

Solution by Peter Borwein, University of British Columbia. In light of 6007, it is sufficient to prove that if f is not singular then $L<2$. Assume f is not singular. Let $H(x)=\int_0^x f'(t)\,dt$, let $G=f-H$ and let $c=H(1)\neq 0$. Define $h=(1/c)H$ and $g=(1-c)^{-1}G$. If $c=1$, then G is 0. Then $f=ch+(1-c)g$, where h and g are nondecreasing continuous functions mapping $[0,1]$ onto $[0,1]$, h is absolutely continuous, g is singular and $c\in(0,1]$. Actually, we may consider only $c\neq 1$, for if $c=1$, the calculation is immediate. Let $0\leqq x_0<x_1<\cdots<x_n\leqq 1$, then

$$
\begin{aligned}
&\sum_{i=1}^{n}[(f(x_i)-f(x_{i-1}))^2+(x_i-x_{i-1})^2]^{\frac12}\\
&=\sum_{i=1}^{n}[(ch(x_i)+(1-c)g(x_i)-ch(x_{i-1})-(1-c)g(x_{i-1}))^2\\
&\qquad+(cx_i+(1-c)x_i-cx_{i-1}-(1-c)x_{i-1})^2]^{\frac12}\\
&\leqq c\sum_{i=1}^{n}[(h(x_i)-h(x_{i-1}))^2+(x_i-x_{i-1})^2]^{\frac12}\\
&\qquad+(1-c)\sum_{i=1}^{n}[(g(x_i)-g(x_{i-1}))^2+(x_i-x_{i-1})^2]^{\frac12}\\
&\leqq cL(h)+(1-c)L(g).
\end{aligned}
$$

Since g is singular, $L(g)=2$ by Problem 6007. Since h is absolutely continuous and $h'>0$ on a set of positive measure, $L(h)=\int_0^1((h')^2+1)^{\frac12}<\int_0^1 h'+1=2$. Thus, $L(f)<2$.

Also solved by R. A. Christiansen, P. R. Goodey (England), M. B. Gregory, Nicholas Passell, and P. van der Steen (Netherlands).

Notes. (1) Van der Steen points out that the methods of solving 6074 (and 6007) also establish the following: If f is real, continuous and of bounded variation, then the arc length of $f(x)$ equals $\int_0^1(1+f'^2)^{\frac12}dx$ + total variation of $f(x)-\int_0^x f'(t)\,dt$.

(2) Roy Davies points out that a more general result has been proved by M. Nevkinda, *Curves of maximum length* (Czech summary), Časopis Pěst. Mat. 99 (1974), 30–43. See Math Reviews, 50, #4859.

(3) W. A. J. Luxemburg points out that the problem (and more) was posed and solved as Problem 5029 [1963, 674] by him and A. C. Zaanen.

(4) Finbarr Holland (Ireland) offers the reference H. Kober, *On singular functions of bounded variation*, J. London Math. Soc. 23 (1948), 222–229.

Integrable Functions with Positive Fourier Transform

6075 [1976, 141]. *Proposed by H. L. Montgomery, University of Michigan*

If $f(x)\in L'(-\infty,\infty)$, $f(x)\geqq 0$, $\hat f(t)\geqq 0$ for all x,t where $\hat f$ is the Fourier transform, prove for any

integer $k \geqq 1$,

$$\int_{-k}^{k} f(x)\,dx \leqq (2k+1)\int_{-1}^{1} f(x)\,dx.$$

Solution by O. P. Lossers, Eindhoven University of Technology, The Netherlands. We shall assume only that $\operatorname{Re}\hat{f}(t) \geqq 0$ for all t, i.e. that

$$\int_{-\infty}^{\infty} \cos tx f(x)\,dx \geqq 0 \qquad (t \in \mathbb{R}). \tag{1}$$

The following identity is readily verified:

$$\begin{aligned} J(a) &= \int_{-a}^{a} (a-|x|)f(x)\,dx \\ &= \frac{a^2}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left\{\frac{\sin(at/2)}{at/2}\right\}^2 \cos tx f(x)\,dt\,dx. \end{aligned} \tag{2}$$

It follows that $\int_{-1}^{1} f(x)\,dx \geqq J(1)$ and that

$$\int_{-k}^{k} f(x)\,dx \leqq J(k+1)-J(k) \ (k=0,1,2,\ldots).$$

Hence it is sufficient to prove that $J(k+1)-J(k) \leqq (2k+1)J(1)$, or using (1) and (2), that

$$\sin^2(k+1)t - \sin^2 kt \leqq (2k+1)\sin^2 t \qquad (k = 0,1,2,\ldots;\ t \in \mathbb{R}).$$

This last inequality is easily verified by using $|\sin kt| \leqq |k \sin t|$ for integer k.

Also solved by Carl Herz, Finbarr Holland (Ireland), A. A. Jagers (Netherlands), Harald Krogstad (Norway), J. Lagarias, A. Odlyzko, and T. Schonbek.

Note. In his solution Krogstad refers to H. S. Shapiro, Quart. J. of Math. 26 (1975), p. 9 ff, for related inequalities.

Sum of the Digits in K^n

6077 [1976, 141] *Proposed by Hugh L. Montgomery, University of Michigan*

Let $s(n)$ denote the sum of the base 10 digits of $(1974)^n$. Show that $s(n) \to \infty$ as $n \to \infty$.

Solution by Michel Mendes France, Université Bordeaux, France. Let $g \geqq 2$ and $h \geqq 2$ be two integers. Let $s_g(n)$ and $s_h(n)$ denote respectively the sum of the digits in base g, respectively base h, of the positive integer n. The following theorem was proved by Senge and Strauss (*PV-numbers and sets of multiplicity.* Periodica Math. Hung. 3, (1973) 93–100):

THEOREM. *Let $A > 1$ and $B > 1$ be two given numbers. Assume that* $\log g/\log h$ *is irrational. Then the set $E = \{n \in \mathbb{N} \mid s_g(n) < A;\ s_h(n) < B\}$ is finite.*

From $s_h(h^n) = 1$ it follows that the set

$$\{n \in \mathbb{N} \mid s_g(h^n) < A\}$$

is finite when $\log g/\log h$ is irrational. In other words

$$\lim_{n\to\infty} s_g(h^n) = +\infty.$$

Choose $g = 10$ and let h be any number which is not a power of 10.

Also solved by H. Niederreiter, and by E. G. Straus. Murray Klamkin observes that the problem has also appeared and been solved in Sierpiński, 250 *Problems in Elementary Number Theory* (pp. 103–104).

Bipartite Graphs

6079 [1976, 205]. *Proposed by D. J. Kleitman, Massachusetts Institute of Technology.*

Given a bipartite graph connecting n vertices with n others. If the symmetry group of the graph is transitive on both parts of the graph, must it be transitive on the whole graph? (Due to Bohdan Zelinka and announced at the Kesthely Conference, June 1973.)

Solution by A.A. Jagers, Technische Hogeschool Twente, Enschede, Netherlands. No. Let G be an edge but not a vertex transitive regular graph. Such graphs exist; see e.g. I.Z. Bouwer, *On edge but not vertex transitive regular graphs*, J. Combinatorial Theory (B) **12**, 32-40 (1972). Then G is bipartite with respect to the two orbits of its symmetry group, and, since G is regular, the two orbits contain the same number of vertices.

Also solved by S.E. Payne, and by the proposer.

Note. Payne's example is a bipartite graph whose vertices are the points and lines of a point-line incidence geometry S, which he constructs and which is a generalized quadrangle of order (q, q) where q is an odd prime power. For the needed properties of S, Payne refers to Benson, *On the structure of generalized quadrangles*, Journal of Algebra, **15** (1970), p. 443 ff.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN
with the assistance of the mathematics departments of St. Olaf and Carleton Colleges
COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, CARLETON COLLEGE

We invite readers to submit reviews of significant recent college-level mathematics books. We especially encourage reviews based on classroom use, or comparative reviews of several related books. Reviews should ordinarily not exceed two pages (per book) typed double spaced. Manuscripts of reviews as well as books submitted for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057.

Euclidean and Non-Euclidean Geometries, Development and History. By Marvin Jay Greenberg. W. H. Freeman and Company, San Francisco, California, 1974. xi + 304 pp. $10.95. (Telegraphic Review, April 1975.)

A college course in geometry is usually required of prospective high school mathematics teachers, presumably to prepare them to teach one of the staples of the secondary mathematics diet. If this purpose is to be served, then the course should bear some similarity to the high school geometry course, both in content and in style. It should expand upon the traditional topics, consider the history of geometry, elucidate its logical and philosophical foundations, and explore one or more generalizations or off-shoots in some detail — all within the synthetic spirit of Euclid. An ideal vehicle is a course in hyperbolic geometry — Euclidean geometry with the negation of the Fifth Postulate. Greenberg's book is an excellent text for such a course. I have used it in a 4-credit semester course for juniors, most of whom were secondary education mathematics majors.

The text's scope is modest and its treatment elementary. There are no formal prerequisites. Using modified versions of Hilbert's axioms for Euclidean geometry minus the parallel postulate, Greenberg develops "neutral geometry" from scratch. The proofs are synthetic and rigorous. Measurement of length and angle is not introduced until nearly the end of this development, at which time a theorem establishing the existence of such measures is stated without proof (here, and throughout, numerous references are provided to enable the reader to fill in missing proofs and delve into topics more

deeply). This basic material occupies Chapters 3 and 4, following an introduction to the flaws of Euclid's original work and an explanation of axiom systems, models, and logic. After discussing equivalents to the parallel postulate and the futile attempts at its proof, Greenberg adopts its negation and develops, in Chapter 6, a bit of the "strange new universe" of hyperbolic geometry. More results are contained in an appendix, and the instructor could easily add to the material presented in the text. Chapter 7 seems slightly more difficult than the previous chapters, as Greenberg proves the consistency of hyperbolic geometry within Euclidean geometry, using the Poincaré disk model and the Beltrami-Klein model. The book concludes with a brief chapter on philosophical implications and has a second appendix touching on other non-Euclidean geometries.

Greenberg's writing style is leisurely and conversational, yet the material is tightly woven. He distinguishes between important "Theorems" and mere "Propositions" and has shown great restraint by omitting interesting but extraneous results which would have interfered with the mathematical flow of the book. The students especially liked the wealth of historical and philosophical material, as well as the twelve full-page portraits.

The exercises at the end of each chapter consist, for the most part, of proving some of the results stated in the chapter or new results needed later. Because of this strong dependence on the exercises, a large fraction of them must be done, either by the student as homework, or by the instructor in lecture. (My students had time to do about half of the exercises as homework.) Some of the problems are difficult, but generous hints and sometimes even outlines of proofs are provided. In addition each chapter ends with about twenty true-false questions, an aid the students liked. There are also some less technical exercises, such as writing an essay on a quotation from G. H. Hardy.

We found few errors, either substantive or typographical. There is some confusion over whether a ray and itself forms an angle of measure zero (it should not), and the statement of Dedekind's Axiom does not allow the cut point to be in either half of the cut.

The students uniformly praised the book for its clarity, readability, enthusiasm, humor, and conciseness. They appreciated being able to complete nearly the entire text in one semester. Said one, "Overall this has been one of the best mathematics textbooks I have used thus far." Given its purpose, I agree.

JERROLD W. GROSSMAN, Oakland University

MISCELLANEA

5. What is a satisfactory definition? For the philosopher or the scholar, a definition is satisfactory if it applies to those and only those things that are being defined; this is what logic demands. But in teaching, this will not do; a definition is satisfactory only if the students understand it.

H. Poincaré, *Science et méthode*, 1909, Book II, Chapter II.

TELEGRAPHIC REVIEWS

Telegraphic reviews are designed to give prompt notice of new books with sufficient information to assist our readers in deciding whether to order an examination copy or to suggest library purchase. Possible uses are indicated as follows:

T = textbook　　P = professional reading
S = supplementary reading　　L = undergraduate library purchase
13 to 18 = freshman to second year graduate level usage
1 to 4 = appropriate time in semesters to cover text

Asterisks (*) or question marks (?) denote special positive or negative emphasis, respectively. Publishers are denoted by standard abbreviations; complete addresses may be found in *Books in Print*.

GENERAL, S*(13-15), L**, *Mathematical Magic Show*. Martin Gardner. Alfred A. Knopf, 1977, 284 pp, $8.95. The eighth collection of Gardner's columns from *Scientific American*, with "reports from readers and afterthoughts from the author." Begins with an informal glossary of mathematical terms, then tours recreational topics from "Nothing" to "Everything" via puzzles and paradoxes, polyominoes and permutations. A sure cure for mathophobia. LAS

GENERAL, T(13-14: 1, 2), *The Business of Mathematics*. Margaret F. Willerding. Prindle, 1977, xii + 639 pp, $12.95. Substantial (5 chapters) review of basic arithmetic and algebra, followed by discussion of topics in business and consumer mathematics, e.g., taxes, annuities, investments, and many others. Chapters on basic statistics, and the metric system included. Problems for hand-held calculators in every chapter. JG

GENERAL, T*(15-17: 1), S*, P, L**, *Encounter with Mathematics*. Lars Gårding. Springer-Verlag, 1977, ix + 270 pp, $9.80. A rich, ambitious exposition of core mathematics written, according to the author, for those in the general public who study mathematics for one year beyond high school. Each chapter moves from elementary to sophisticated results (e.g., Roth's 1955 bound for the Dirichlet approximation theorem, Hilbert's Nullstellensatz, the spectral theorem for self-adjoint operators, the deRham complex, and cohomology groups, Fourier inversion formula, Weierstrass approximation theorem), and concludes with brief extracts from relevant source documents. Much more sophisticated than Courant and Robbins' classic *What is Mathematics?*, *Encounter* is, realistically, a unique synopsis of classical mathematics best suited to upperclass mathematics majors. LAS

GENERAL, T(13: 1, 2), *Mathematics as a Second Language, Second Edition*. Frances Lake, Joseph Newmark. A-W, 1977, xviii + 491 pp, $13.95. The exposition remains suitable for students with little or no high school mathematics. A chapter on computers and elementary Fortran is new, and the chapters on geometry, probability and statistics and number bases have been broadened to respond to the needs of students in elementary education or business. (First edition, TR, May 1974.) JG

GENERAL, S?(13), *Mathematics Unraveled--A New Commonsense Approach*. James Kyle. TAB Books, 1976, 279 pp, $9.95; $6.95 (P). Mathematics for the man-in-the-street. Generalities and homilies interspersed throughout a hodgepodge collection of facts and rules drawn from set theory, logic, arithmetic, algebra, calculus, probability and statistics. LCL

GENERAL, S(15), *Mathematical Formulas for Engineering and Science Students, Second Edition*. S. Barnett, T.M. Cronin. Bradford U Pr, 1975, viii + 74 pp, £1.95 (P). Lists of formulas pertaining to series, elementary transcendental functions, indefinite integrals, integration methods, ordinary differential equations, Laplace transforms, vector analysis, matrices and determinants, numerical methods, statistics, lengths, areas, volumes, moments. Tables. Selected for engineering students. DFA

GENERAL, P, *Lecture Notes in Mathematics-567: Séminaire Bourbaki vol. 1975-76, Exposés 471-488*. Springer-Verlag, 1977, iv + 303 pp, $13.70 (P).

GENERAL, S*, *Guidelines for the Tutor of Mathematics*. Henry S.Kepner, Jr., David R. Johnson. NCTM, 1977, iii + 28 pp, $1.30 (P). Excellent, useful pamphlet with practical ideas and suggestions for effective tutoring. Here is a source that can be read in a single sitting and reread throughout the tutor's experience, and can serve as a basis for regular training and discussion sessions. Discounts on quantity orders. LCL

GENERAL, S*, *How to Study Mathematics*. James Margenau, Michael Sentlowitz. NCTM, 1977, ii + 31 pp, $1.30 (P). Helpful, concrete suggestions and prescriptions for common complaints written in a warm, personal, easy-to-read manner. Several copies should be in every student academic support center. Discounts on quantity orders. LCL

GENERAL, T*(13-14: 1, 2), *Fundamental Mathematics, A Cultural Approach*. A. Richard Polis, Earl M.L. Beard. Har-Row, 1977, xxi + 538 pp, $14.95. Conversational without being condescending. Material covered includes chapters developing the real number system, algebraic structures, geometry, probability and statistics. Special feature: thirty-two short "capsules" of an historical or cultural nature with bibliographies. Good problems; self-tests for each chapter. JG

GENERAL, S(9-13), L**, *Asimov on Numbers*. Isaac Asimov. Doubleday, 1977, xv + 249 pp, $9.95. Seventeen popular essays on numbers and their relation to measurements, science, calculation and mathematics. Each originally appeared between 1959 and 1966 in *The Magazine of Fantasy and Science Fiction*. LAS

GENERAL, T*, S**(13), L**, *The Method of Mathematical Induction*. I.S. Sominsky. Trans: Martin Greendlinger. MIR (Imported by: Imported Pub, 320 W. Ohio St., Chicago, IL 60610), 1975, 62 pp, $1 (P). Beautiful exposition in a classic style of an important topic. Careful development explores, by examples drawn from algebra and trigonometry, why induction works, when it can be expected to apply, and how it is applied. Bargain price makes it especially suitable as a course supplement. LCL

BASIC, T(13), S*. *The Metric System, Module SI, Series in Mathematics Modules.* Judith L. Gersting, Elaine V. Alton. Cummings, 1977, 124 pp, $3.95 (P). Seven self-study modules covering length (together with applications), mass, and temperature. LCL

PRECALCULUS, T(13: 1). *Plane Trigonometry.* Michael E. Bennett, Richard A. Miller, Barry N. Stein. Saunders, 1977, xvi + 430 pp, $11.50. Classical approach to trigonometry. Includes complex numbers and vectors. End of each chapter has a summary, review exercises and practice exam. LLK

PRECALCULUS, T(13: 1). *Functions and Graphs.* Earl W. Swokowski. Prindle, 1977, vii + 496 pp, $13.95. This is a good precalculus text (not an algebra text) on functions (polynomial, rational exponential, logarithmic, and trigonometric). LLK

PRECALCULUS, T*(13: 1). *Fundamental Algebra and Trigonometry.* Mervin L. Keedy, Marvin L. Bittinger. A-W, 1977, xiii + 498 pp, $13.95; *Study Supplement,* 284 pp, (P). Clear explanations and format. Many diagrams and worked-out examples. An abundance of representative exercises. Exercises for electronic calculators. Chapter tests. Answers to selected problems. Tables. Index. Also available: *Answer Booklet, Diagnostic Test Booklet.* RJA

PRECALCULUS, T(13: 1). *Trigonometry.* Margaret L. Lial, Charles D. Miller. Scott F, 1977, 312 pp, $11.95. Uses the metric system throughout. Many examples worked out and numerous exercises with answers to odd-numbered problems. Chapter summaries and tests. Format of the text is very good. Some problems designed for pocket calculators. Tables. *Solutions Guide, Instructor's Guide and Mathlab, Study Guide* all available. Index. RJA

EDUCATION, T(14: 1), S. *Elementary School Mathematics for Teachers.* C. Alan Riedesel, Leroy G. Callahan. Har-Row, 1977, vii + 311 pp, $8.95 (P). A modularized approach to mathematics for the elementary school teacher, complete with objectives, pre-tests, post-tests, and lab exercises. A reasonable approach to a course in which student background varies greatly. CEC

HISTORY, S(17-18), P, L. *Functional Analysis in Historical Perspective.* A.F. Monna. Oosthoek Pub, 1973, viii + 167 pp, $12 (P). Paperback edition of hardcover original (TR, May 1974; ER, January 1977). LAS

NUMBER THEORY, P. *Algebraic Number Fields (L-Functions and Galois Properties).* Ed: A. Fröhlich. Acad Pr, 1977, xii + 704 pp, $31.25. Proceedings of a 1975 conference held in Durham, England. Contains 20 papers including long articles by Martinet (nonabelian L-functions), Tate (root numbers), Fröhlich (Galois module structure of rings of integers), Serre (modular forms and Galois representations), and Coates (p-adic L-functions and Iwasawa's theory). SG

LINEAR ALGEBRA, T(14: 1), S, L. *Applications of Linear Algebra.* Chris Rorres, Howard Anton. Wiley, 1977, ix + 233 pp, $4.95 (P). Applications include Markov chains, forest management, applications in genetics, and a 3-chapter mini-course in linear programming. Chapters are more or less independent, and of varying difficulty, as rated (accurately) by the authors. Exercises for each chapter. Good supplement for standard sophomore-level course. JG

ALGEBRA, T**(15-17: 1, 2), S, L*. *Algebra: Groups, Rings, and Other Topics.* Neal H. McCoy, Thomas R. Berger. Allyn, 1977, xiii + 658 pp, $17.95. Considerable overlap with McCoy's *Fundamentals of Abstract Algebra* (TR, October 1972; ER, August-September 1974). A complete rewriting of the introductory chapter on group theory, which is now placed before the chapter on rings, followed later by more group theory, provides a substantial increase in coverage. Commentaries added at the end of each chapter feature historical notes, suggestions for readings, essays on mathematics, etc. LCL

ALGEBRA, P. *Skew Field Constructions.* P.M. Cohn. London Math. Soc. Lect. Notes, No. 27. Cambridge U Pr, 1977, xii + 253 pp, $12.95 (P). Focuses on the coproduct construction which produces, for every pair of skew fields, another skew field. The author also discusses skew field extensions, polynomial and rational identities, and equations over skew fields. A valuable reference for algebraists. SG

ALGEBRA, P. *Lecture Notes in Mathematics-573: Group Theory.* Ed: R.A. Bryce, J. Cossey, M.F. Newman. Springer-Verlag, 1977, 146 pp, $8 (P). Proceedings of the conference held at the Australian National University, Canberra, November 4-6, 1975. JAS

ALGEBRA, P. *Logique des Topos (Introduction à la Théorie des Topos Élémentaires).* Dana I. Schlomiuk. Pr U Montreal, 1977, 132 pp, $5 (P). Notes from a June 1974 course entitled "Méthodes catégoriques en logique mathématique et théorie des automates." JAS

ALGEBRA, T(17-18: 1, 2), S. *Moduln und Ringe.* Friedrich Kasch. Teubner, Stuttgart, 1977, 328 pp, DM 52 (P). A text on modules and rings. Treats first the basic theory and then specialized results on rings with complete duality and quasi-Frobenius rings. Exercises and a limited bibliography. JD-B

ALGEBRA, P. *The Structure of Modular Lattices of Width Four with Applications to Varieties of Lattices.* Ralph S. Freese. Memoirs No. 181. AMS, 1977, vii + 91 pp, $7.20 (P). The lattice variety generated by all modular lattices of width not exceeding four is shown to be finitely based, and the method yields a complete list of subdirectly irreducible width four modular lattices. Appendices include earlier related results. CEC

ALGEBRA, T(17-18: 1, 2), P. *Lectures in Semigroups.* Mario Petrich. Akademie-Verlag, 1977, viii + 168 pp. Written to serve as a treatise on the theory of semigroups, or as an advanced text, it concentrates on bands, matrix and normal band decompositions, and lattices of subsemigroups. Exercises and a considerable bibliography. JD-B

ALGEBRA, P. *Ring Theory.* Ed: S.K. Jain, Klaus E. Eldridge. Lect. Notes in Pure and Appl. Math., V. 25. Dekker, 1977, viii + 246 pp, $24.50 (P). The invited papers from a conference at Ohio University in May 1976. Papers are mainly in algebraic geometry and topological algebra. JAS

FINITE MATHEMATICS, T**(13-16: 1), L. *Applied Finite Mathematics*. Robert F. Brown, Brenda W. Brown. Wadsworth, 1977, x + 549 pp, $13.95. A unique fleshed-out, finite-math book. The usual topics are here, but they are treated in context. Each chapter has a specific real-life social science or business situation as a theme, e.g., detection of underground nuclear tests, farm management, Prisoner's Dilemma, opinion sampling, home mortgages. Each chapter begins with an "introductory essay" giving the setting and a problem, and ends with a "concluding essay" returning from the mathematics to the original situation. Most chapters also contain either a "technical essay" or a "historical essay", and references to the sources for the data. PJC

CALCULUS, S, P. *Aufgabensammlung zur Infinitesimalrechnung, Band III: Integralrechnung auf dem Gebiete mehrerer Variablen*. A. Ostrowski. Math. Reihe, B. 56. Birkhäuser, 1977, 398 pp, sFr. 68. Third and final volume of a fine collection of problems which parallels the author's calculus text. Deals with integral calculus, chiefly of functions of several variables. Problems, hints and solutions printed separately. JD-B

CALCULUS, S(13). *Programmed Guide to Accompany Calculus by Howard E. Campbell & Paul F. Dierker*. Roy A. Dobyns. Prindle, 1977, vii + 236 pp, $4.95 (P). Programmed supplement for the first 8 chapters (there are 16 chapters) of *Calculus* by Campbell and Dierker (TR, May 1975). LLK

REAL ANALYSIS, S(16-17), P, L. *Orthogonal Functions, Revised English Edition*. G. Sansone. Trans: Ainsley H. Diamond. Krieger, 1977, xii + 411 pp, $19.50. A reprinting of the 1959 Wiley revised English translation of the 1952 third edition of Sansone's 1935 sequel to G. Vitali's *Moderna Teoria delle Funzioni d. Variabili Reale*. Hilbert space, Fourier, Legendre, Laguerre and Hermite series. LAS

COMPLEX ANALYSIS, P. *Several Complex Variables*. Ed: R.O. Wells, Jr. Proc. of Symp. in Pure Math., V. XXX. AMS, 1977. *Part 1*, x + 390 pp; *Part 2*, xii + 328 pp, $44.40. Principal lectures, hour lectures and research reports, arranged in eight seminar series. Volume I contains series on singularities of analytic spaces, function theory, and compact complex manifolds. Volume II contains series on noncompact complex manifolds, differential geometry, approximation, value distribution theory, and harmonic analysis. LAS

COMPLEX ANALYSIS, P. *Théorie des Fonctions Algébriques, Tome 1: Étude des Fonctions Analytiques sur une Surface de Riemann*. Paul Appell, Édouard Goursat, P. Fatou. Chelsea, 1976, xxxv + 526 pp, $29.50. Reprint of the second edition of a work first published in 1929. A clear and thorough treatment of function theory on Riemann surfaces. Topics include: hyperelliptic integrals, periodicity of abelian integrals, Riemann-Roch, and Abel's theorem. SG

DIFFERENTIAL EQUATIONS, L. *Elementary Differential Equations, Third Edition*. William E. Boyce, Richard C. DiPrima. Wiley, 1977, xiv + 496 pp, $15.95. The major changes included in this edition are an extensive review of power series, applications of systems of equations to network analysis, emphasis on the Adams-Moulton method, more material on stability theory, and 150 more problems. (TR, *Second Edition*, October 1969 and March 1977; ER, October 1977.) SG

DIFFERENTIAL EQUATIONS, T(14-15), L. *An Introduction to Ordinary Differential Equations with Difference Equations, Numerical Methods, and Applications*. Garret J. Etgen, William L. Morris. Har-Row, 1977, x + 517 pp, $14.95. A comprehensive survey for undergraduates. A special feature is the inclusion of difference equations. Also covered are linear equations (including linear systems) numerical methods, series solutions, Laplace transforms and a variety of applications. Many routine and non-routine exercises. SG

DIFFERENTIAL EQUATIONS, P. *Differential Equations*. Ed: M. Farkas. North-Holland, 1977, 418 pp, $48. Proceedings of a 1974 Hungarian conference. The 33 pages range over such topics as oscillation, stability, delay-differential equations, existence and uniqueness, numerical solutions, partial differential equations, and Hamiltonian systems. SG

DIFFERENTIAL EQUATIONS, P. *Theory of Functional Differential Equations*. Jack Hale. Appl. Math. Sci., V. 3. Springer-Verlag, 1977, x + 365 pp, $24.80. An expansion of the author's *Functional Equations*. Among the new topics are linear differential difference equations, dissipative systems, perturbed systems, periodic solutions of autonomous equations. SG

DIFFERENTIAL EQUATIONS, P. *Stability of Dynamical Systems, Theory and Applications*. Ed: John R. Graef. Lect. Notes in Pure and Appl. Math., V. 28. Dekker, 1977, xi + 214 pp, $19.75 (P). Proceedings of the Regional Research Conference sponsored by NSF and CBMS at Mississippi State University in August 1975. The principal lectures, by J.P. LaSalle, appear as Volume 25 in the SIAM Regional Conference Series in Applied Mathematics. JAS

NUMERICAL ANALYSIS, P. *Fixed Points, Algorithms and Applications*. Ed: Stepan Karamardian. Acad Pr, 1977, x + 494 pp, $19.50. The proceedings of the International Conference on Computing Fixed Points with Applications which was held at Clemson University in June 1974. Herb Scarf has provided an introductory essay. JAS

NUMERICAL ANALYSIS, S(17-18), P. *Quadraturverfahren*. Helmut Brass. Vandenhoeck & Ruprecht, 1977, 311 pp, DM 35 (P). On the theory of various methods of approximating the definite integral of continuous functions. Devoted largely to work done since 1960. JD-B

NUMERICAL ANALYSIS, T(18: 1), S, P. *Construction of Integration Formulas for Initial Value Problems*. P.J. van der Houwen. Appl. Math. and Mech., V. 19. North-Holland, 1977, xi + 269 pp, $33.95. Single and multi-step methods. Classified as Taylor's, Runge-Kutta, and "generalized" Runge-Kutta. Uses a control function to obtain adaptive methods, in particular, for stiff problems and for partial differential equations. Includes a study of stability polynomials. RWN

NUMERICAL ANALYSIS, P, *Factorization Methods for Discrete Sequential Estimation.* Gerald J. Bierman. Math. in Sci. and Eng., V. 128. Acad Pr, 1977, xvi + 241 pp, $19. A development of the computational aspects of the efficient square root alternatives to the Kalman filter for discrete linear estimation problems. Includes some preliminary computational linear algebra and programs. RWN

NUMERICAL ANALYSIS, P, *Lecture Notes in Mathematics-572: Sparse Matrix Techniques.* Ed: V.A. Barker. Springer-Verlag, 1977, 184 pp, $8 (P). Notes from the advanced course held at the Technical University of Denmark, Copenhagen, on August 9-12, 1976. JAS

ANALYSIS, P, *Littlewood-Paley and Multiplier Theory.* R.E. Edwards, G.I. Gaudry. Ergebnisse der Math., B. 90. Springer-Verlag, 1977, ix + 212 pp, $25.60. An introduction. Various versions of the Littlewood-Paley and Marcinkiewicz multiplier theorems for T^n, Z^n, and R^n, and for certain totally disconnected groups (including a martingale version for these). Applications to lacunary sets and Fourier multiplier theory. DFA

ANALYSIS, T(15: 1), *Vector Analysis: A Physicists' Guide to the Mathematics of Fields in Three Dimensions.* N. Kemmer. Cambridge U Pr, 1977, xiv + 254 pp, $28.50; $8.95 (P). Presentation based on a general parametrization of curves and surfaces. Treats standard topics, but with emphasis on visualization. Attractively illustrated. Many intriguing exercises; hints and answers. From a course at Edinburgh. DFA

ANALYSIS, P, *The Theory of Ultraspherical Multipliers.* W.C. Connett, A.L. Schwartz. Memoirs No. 183. AMS, 1977, iv + 92 pp, $7.20 (P). An organization of the multiplier theorems for ultraspherical expansions together with many new results. JAS

ALGEBRAIC GEOMETRY, P, *Lecture Notes in Mathematics-585: Invariant Theory.* T.A. Springer. Springer-Verlag, 1977, 112 pp, $8 (P). An "elementary level" introduction. Focuses on the notion of a (geometrically) reductive linear algebraic group. For such a group the ring of invariant polynomials is of finite type. The author proves that $Sl_2(k)$ is reductive and discusses invariant theory for finite reflection groups. Interesting historical notes are included. Useful to algebraists, geometers and coding theorists. SG

ALGEBRAIC GEOMETRY, P, *Conference on Quadratic Forms-1976.* Ed: G. Orzech. Pure and Appl. Math., No. 46. Queen's U, 1977, iii + 656 pp, (P). Proceedings of the conference which took place in August 1976 at Queen's University. JAS

ALGEBRAIC GEOMETRY, P, *Lecture Notes in Mathematics-569: Cohomologie Etale.* P. Deligne. Springer-Verlag, 1977, v + 312 pp, $13.70 (P). A special "expository" volume on ℓ-adic cohomology developed from SGA4 (Lecture Notes 269, 270 and 305) and serving as an introduction to SGA5 which is still in preparation. JAS

DIFFERENTIAL GEOMETRY, S(17-18), *The Absolute Differential Calculus (Calculus of Tensors).* Tullio Levi-Civita. Dover, 1977, xvi + 452 pp, $6 (P). An unaltered and unabridged republication of the English translation (Blackie and Son, Ltd., 1926) of the original Italian edition. JAS

DIFFERENTIAL GEOMETRY, P, *Geometric Asymptotics.* Victor Guillemin, Shlomo Sternberg. Math. Surveys, No. 14. AMS, 1977, xviii + 474 pp, $34.40. Using the language of differential geometry, the authors develop the themes of symplectic geometry and Fourier integral operators, two subjects rooted in the relations between the wave and corpuscular theories of light. Topics: method of stationary phase, differential operators and asymptotic solutions, compound asymptotics, optics, quantization, and distributions. TRS

TOPOLOGY, P, *Set-Theoretic Topology.* Ed: George M. Reed. Acad Pr, 1977, xv + 436 pp, $18. Papers from the conferences during the 1975-1976 academic year at the Institute for Medicine and Mathematics at Ohio University. Emphasis is on the recent developments in set theory and their interaction with point set topology. JAS

TOPOLOGY, S(18), P, *Foundational Essays on Topological Manifolds, Smoothings, and Triangulations.* Robion C. Kirby, Laurence C. Siebenmann. Annals of Math. Stud., No. 88. Princeton U Pr, 1977, ix + 355 pp, $19.50; $8.50 (P). A consolidation of recent gains in the theory of topological manifolds offering both a survey of topological manifolds and of classification of smoothings and triangulations. Not elementary, but written with a concern for the reader. JAS

PROBABILITY, T(13-14: 1), *A Programmed Introduction to the Theory of Probability.* Dieter Stempell. Verlag Die Wirtschaft, 1973, 168 pp, $5.50 (P). Formal approach beginning with laws of probability, continuing through binomial, Poisson, normal, exponential and beta distributions. No proofs. Weak on motivation and application. Intended for beginners, yet presumes familarity with elementary combinatorics, limits and calculus. Translated from German edition. LCL

PROBABILITY, P, *Stochastic Integration and Generalized Martingales.* A.U. Kussmaul. Pitman, 1977, 163 pp, £7 (P). The author attempts to provide a unified approach to stochastic integration directly through the properties of stochastic processes, to arrive at the integrals of Weiner, Ito and Doob through probabilistic arguments. Includes a discussion of generalized martingales. TAV

PROBABILITY, P, *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations.* Harold J. Kushner. Math. in Sci. and Eng., V. 129. Acad Pr, 1977, xvii + 243 pp, $23.

STATISTICS, P*, *Sufficient Statistics, Selected Contributions.* Vasant S. Huzurbazar. Statistics, V. 19. Dekker, 1976, xiii + 270 pp, $27.50. Collection of the author's rigorous research monographs. Part 1 deals with Bayesian inference and invariance theory of prior probabilities, primarily developing a theory of new invariants of distributions admitting sufficient statistics. Part 2 is on general forms of distributions admitting sufficient statistics for parameters in nonregular cases, while Part 3 deals with location and scale parameters and sufficient statistics. RSK

STATISTICS, T(13-14: 1). *Elementary Statistical Methods in Psychology and Education, Second Edition.* Paul J. Blommers, Robert A. Forsyth. HM, 1977, xxii + 570 pp, $13.95; *Study Manual,* vii + 258 pp, $6.25 (P). Extensive coverage is sacrificed for an unusually thorough discussion of a small number of basic concepts and techniques (descriptive statistics, estimation and hypothesis testing for means and proportions, correlation, regression). New edition includes optional chapters on probability and Bayesian inference. Not included: chi-squared and F-distributions, analysis of variance, nonparametrics. No exercises in text. Excellent workbook (must be read with text) contains typical educational settings followed by leading questions and exercises. Answers not provided. LCL

STATISTICS, S(17), L*. *Mathematical Methods for Digital Computers, V. III: Statistical Methods for Digital Computers.* Ed: Kurt Enslein, Anthony Ralston, Herbert S. Wilf. Wiley, 1977, viii + 454 pp, $24.95. Continues the excellent quality of the first two volumes. 15 survey articles by 20 authorities covering regression, discriminant analysis, factor analysis, multivariate analysis of variance, cluster analysis, the fast Fourier transform, and time series. Includes flow charts and a list of program sources. RWN

STATISTICS, T??(13: 1). *Introduction to Statistics, Revised Edition.* Robert Fried. Gardner Pr, 1976, xvi + 304 pp, $12.95. Incredibly confusing text, filled with misinformation throughout. For example, there are only two sentences concerning conditional probability and one states, "If the events E_1 and E_2 are *not* mutually exclusive, e.g., not independent, their probability is the product of their respective conditional probabilities." RSK

STATISTICS, T*(13), S. *Understanding Statistics, Second Edition.* Arnold Naiman, Robert Rosenfeld, Gene Zirkel. McGraw, 1977, xii + 307 pp, $11.95. Presumes little in mathematical prerequisites. Good intuitive discussion of probability, binomial, normal distributions, and binomial approximations to normal. Hypothesis testing of proportions (one, two sample), and means (large samples, small samples). Confidence intervals for these cases. Second edition adds to chapters on confidence intervals and chi square tests, and adds eight chapters on analysis of variance and nonparametric methods. Examples and exercises entertainingly chosen. (First edition, TR, February 1973.) RBK

STATISTICS, T(14: 1, 2). *Statistics for Management Decisions.* Donald R. Plane, Edward B. Oppermann. Business Pub, 1977, xv + 527 pp, $14.95. For students of business and economics. Presumes only high school algebra, but requires some sophistication. Views both estimation and hypothesis testing as information models rather than decision procedures. Has two chapters on decision analysis, and chapters on multiple regression, time series analysis, and index numbers, in addition to standard topics. Last quarter of the book contains mathematical notes for the student, tables, and selected answers. RSK

STATISTICS, T*(15-17: 2, 3), L. *Statistics and Experimental Design in Engineering and the Physical Sciences, Second Edition.* Norman L. Johnson, Fred C. Leone. Wiley, 1977. *V. I,* xiv + 601 pp; *V. II,* xiv + 498 pp, $24.95 each. Revision, incorporating a substantial number of minor changes, of the first (1964) edition. Volume I contains the usual major topics plus chapters on order statistics, distribution-free methods, control charts, and decision theory. About two-thirds of Volume II is devoted to analysis of variance, with other chapters covering sequential analysis, multivariate analysis, response surfaces, and sampling techniques. A nice blend of theoretical and practical material. A significant change is the increased number of solutions to exercises. RSK

STATISTICS, T(13: 1, 2). *Statistical Reasoning in Sociology, Third Edition.* John H. Mueller, Karl F. Schuessler, Herbert L. Costner. HM, 1977, xvi + 544 pp, $13.95. Extensive coverage of descriptive statistical techniques (2/3 of the book), together with major topics of inferential statistics. New topics in this edition include component analysis, multiple correlation and regression, path analysis, analysis of variance, and multivariate contingency table analysis, all of which are more technical and may be postponed to a second term or course. RSK

STATISTICS, T*(17), P. *Linear Regression Analysis.* G.A.F. Seber. Wiley, 1977, xvii + 465 pp, $29.95. In the Wiley Series in Probability and Mathematical Statistics. A theoretical text which also provides up-to-date material on computational methods and algorithms. Emphasizes basic results but also provides information about more general procedures. Good problem sets, with outline solutions provided, and a good set of references. RSK

STATISTICS, T(18), P*. *Methods for Statistical Data Analysis of Multivariate Observations.* G. Gnanadesikan. Wiley, 1977, x + 311 pp, $19.95. In the Wiley Series in Probability and Mathematical Statistics. Concerned with methodology, rather than theory, in five general areas: reduction of dimensionality, development and study of multivariate dependencies, classification and clustering, assessment of specific aspects of models, and summarization and exposure. Results are mostly recent and are illustrated with a wide variety of graphical techniques. Good set of references. Could be used as a text with local data sets for problems. RSK

STATISTICS, P*. *Statistical Decision Theory and Related Topics II.* Ed: Shanti S. Gupta, David S. Moore. Acad Pr, 1977, xiii + 478 pp, $19.50. Contains the 26 invited papers presented at an international symposium on Statistical Decision Theory and Related Topics held at Purdue University in May, 1976. (First symposium papers TR, April 1972.) Emphasizes general decision theory, multiple decision theory, optimal experimental design, and robustness. RSK

STATISTICS, P. *Multivariate Analysis in Behavioral Research.* A.E. Maxwell. Chapman & Hall, 1977, ix + 164 pp, $8.95 (P). One of the Halsted Press (formerly Methuen's) Monographs on Applied Probability and Statistics. Concise, introductory presentation of classical techniques, presuming a working knowledge of matrix algebra (presented in Chapter 3). Concludes with a chapter by B.S. Everitt on cluster analysis and other exploratory techniques. RSK

STATISTICS, T(13: 1), S. *Elementary Probability and Statistical Reasoning*. Howard E. Reinhardt, Don O. Loftsgaarden. Heath, 1977, xiii + 417 pp, $12.95. A noncalculus presentation which emphasizes simulation and model building. Makes excellent use of tree diagrams to understand and calculate probabilities. Interesting problems. RBK

STATISTICS, T(13-14: 1, 2). *Fundamental Statistics for Business and Economics*. Thomas R. Dyckman, L. Joseph Thomas. P-H, 1977, xv + 715 pp, $16.95. Sound and sprightly introduction to probability and statistics at noncalculus level for business students. Problems emphasize managerial decisions. Each chapter ends with a warning, pointing out limitations. RBK

STATISTICS, T*(14-15: 1, 2). *Probability and Statistics for Engineers, Second Edition*. Irwin Miller, John E. Freund. P-H, 1977, xii + 529 pp, $16.50. Updated version of the authors' 1965 first edition. Chapter on applications to operations research has been woven into other chapters; remaining chapter topics are the same, including factorial experimentation, applications to quality assurance, and applications to reliability and life testing. Analysis of covariance has been added and treatment of operating characteristic curves expanded. Requires some calculus, but is a very readable text. RSK

STATISTICS, T(13: 1). *Introductory Statistics with Applications*. Wayne W. Daniel. HM, 1977, xii + 475 pp, $13.95. Noncalculus approach includes probability, estimation, hypothesis testing, analysis of variance, regression, nonparametrics. Straightforward, readable text with emphasis on practicality and intuition. Lots of exercises from a wide variety of realistic settings. LCL

STATISTICS, T(13: 1). *Introductory Biostatistics for the Health Sciences*. Robert C. Duncan, Rebecca G. Knapp, M. Clinton Miller III. Wiley, 1977, vii + 163 pp, $8.50 (P). Written and used as a self-paced text for first year medical students and nurses, with examples and problems drawn from common medical situations. Bare bones coverage, however includes useful analysis and interpretative remarks. Omits inference concerning binomial parameters, as well as nonparametrics. LCL

STATISTICS, T(13: 1). *Statistics Step by Step*. Howard B. Christensen. HM, 1977, xvi + 669 pp, $13.95 (P). Quite complete coverage carefully laid out in module form for self study--twelve chapters broken into 92 lessons systematically structured into objectives, definitions, discussion, examples, and problems. Easy to follow. Useful *Instructor's Manual*. LCL.

STATISTICS, P*. *The Theory and Applications of Reliability, With Emphasis on Bayesian and Nonparametric Methods*. Ed: Chris P. Tsokos, I.N. Shimi. Acad Pr, 1977. *V. I*, xvi + 549 pp, $25; *V. II*, xvi + 582 pp, $26. Proceedings of a Conference held at the University of South Florida in December, 1975. Volume I contains the theoretical presentations, while Volume II consists of the more practical papers. Both contain a survey of recent work in the area by the editors. RSK .

STATISTICS, T(14-16: 1, 2). *Applied Statistics for Science and Industry*. Albert Romano. Allyn, 1977, xiv + 513 pp, $17.95. Designed for a second course in statistics at the pre-calculus level. Chapters 2-5 provide a review of estimation, hypothesis testing, regression and correlation, but at a more sophisticated level, and also include some additional topics, such as Bayesian inference. Chapters 6-7 discuss goodness of fit and other nonparametric tests, while Chapters 8-11 deal with analysis of variance and analysis of covariance. Contains many realistic examples and exercises. RSK

STATISTICS, T(13-14: 1, 2). *Basic Statistics for Business and Economics, Second Edition*. Paul G. Hoel, Raymond J. Jessen. Wiley, 1977, xii + 536 pp, $14.95. Same topics as first edition (TR, March 1973) with expanded coverage, 50% more illustrations, examples and problems. LCL

STATISTICS, T(13: 1). *Introduction to Statistical Methods*. Basil P. Korin. Winthrop Pub, 1977, xvi + 426 pp, $13.95. Emphasizes ordinal and nominal data, and hence nonparametric techniques, more than most texts at this level. Main concern is with statistical inference, but considerable attention is given to descriptive techniques. Treatment of probability distributions and random variables is terse. RSK

STATISTICS, P. *Empirical Distributions and Rank Statistics*. M.C.A. van Zuijlen. Math. Centre Tracts, No. 79. Math Centrum, 1977, vii + 92 pp, Dfl. 12 (P). Monograph proving asymptotic normality of linear rank statistics in cases where sample elements are independent but not identically distributed. Uses new properties of empirical distribution functions, which are first derived. RSK

STATISTICS, T*(13: 1). *Elementary Statistics*. John A. Ingram. Cummings, 1977, xi + 445 pp, $13.95. Very readable introduction in an attractive format. More elementary than the author's 1974 text *Introductory Statistics* (TR, May 1976)--no nonparametric tests, and other topics are limited to basic concepts. Contains seven good essay examples illustrating practical uses of statistics. Appendix contains a mathematics review. RSK

STATISTICS, T*(13: 1). *Elements of Statistical Inference, Fourth Edition*. David V. Huntsberger, Patrick Billingsley. Allyn, 1977, ix + 385 pp, $14.95. Modest revision, including revised problem sets and a more attractive format, of the authors' 1973 *Third Edition* (TR, March 1974). A sound elementary text. RSK

STATISTICS, P*. *Case Studies in Sample Design*. A.C. Rosander. Statistics, V. 21. Dekker, 1977, viii + 426 pp, $35. Down-to-earth book on the application of the theory of probability sampling to specific management problems. First third deals with foundations, but is primarily practical in its orientation. Last two-thirds presents case studies, 15 in capsule form and 12 in great detail. Covers all six phases of a sample study: planning, design, implementation, processing, interpretation and appraisal. RSK

STATISTICS, S, P*. *Pocket Book of Statistical Tables*. Robert Odeh, *et al.* Statistics, V. 22. Dekker, 1977, x + 166 pp, $8.75 (P). Compact collection of modern statistical tables, eliminating tables made obsolete by hand calculators. Emphasizes tables for nonparametric tests. RSK

STATISTICS, T(14-16: 1, 2). *Statistical Concepts and Methods.* Gouri K. Bhattacharyya, Richard A. Johnson. Wiley, 1977, xv + 639 pp, $14.95. In the Wiley Series in Probability and Mathematical Statistics. Provides extensive coverage of the standard topics, as well as some nonstandard ones (e.g., sample surveys), in most cases going beyond what is usually covered in a precalculus text. Emphasizes assumptions and contains more than the usual number of real examples and exercises. RSK

STATISTICS, S?(13). *Everything You Always Wanted To Know About Elementary Statistics (but were afraid to ask).* Jerald G. Schutte. P-H, 1977, x + 230 pp, $6.95 (P). Interview format with questions coming from social science students trapped in a beginning statistics course. The questions span the full spectrum of topics, but the presentation is appropriate only as a supplementary text. Answers, emphasizing straight talk, analogy and metaphor, vary widely in quality. Material on slide rule is unnecessary. LCL

STATISTICS, T*(1). *Intermediate Business Statistics, Analysis of Variance, Regression, and Time Series.* Robert B. Miller, Dean W. Wichern. HR&W, 1977, xiv + 525 pp, $20. In the Series in Quantitative Methods for Decision Making. Presumes a first course in statistics and some exposure to calculus and linear algebra, all of which are reviewed. Concentrates on fixed effect analysis of variance models and basic results in regression and time series analysis, illustrated with real data. RSK

STATISTICS, T(17), P*. *Discrete Multivariate Analysis: Theory and Practice.* Yvonne M.M. Bishop, *et al.* MIT Pr, 1977, x + 557 pp, $15 (P). Paperback edition of the authors' definitive 1975 book (TR, December 1975). RSK

COMPUTER SCIENCE, T(13-16: 1), S. *Programmierung mit Fortran.* Wolfgang Brauch. Teubner, Stuttgart, 1977, 189 pp, DM 10,80 (P).

COMPUTER SCIENCE, T(17: 1), S, P*, L. *Compiler Construction, An Advanced Course, Second Edition.* Ed: F.L. Bauer, J. Eickel. Springer-Verlag, 1976, xiv + 638 pp, $10.60 (P). Based on a short course offered in 1974 and repeated in 1975. Contributions by Bauer, DeRemer, Ershov, Gries, Griffiths, Hill, Horning, Koster, McKeeman, Poole and Waite. A broad survey useful to anyone with an interest in and some knowledge of compilers. Good bibliographies. RWN

COMPUTER SCIENCE, P. *Lecture Notes in Computer Science-42: Complementary Definitions of Programming Language Semantics.* James E. Donahue. Springer-Verlag, 1976, 172 pp, $8.20 (P). Demonstrates on a Pascal subset the utility and limitations of defining semantics at two levels using a mathematical approach and an axiomatic approach to provide the basics for proving the consistency of the definitions and the correctness of implementations. RWN

COMPUTER SCIENCE, P. *Associons and the Closure Statement.* M. Rem. Math. Centre Tracts, No. 76. Math Centrum, 1976, vii + 115 pp, Dfl. 14 (P). A monograph on an extension of programming language theory to effectively account for concurrency of elementary operations with an associative memory. Closure statements play a central role. Sample programs are analyzed. Bibliography. RWN

COMPUTER SCIENCE, T(15: 1), P. *Digital System Design Automation: Languages, Simulation & Data Base.* Ed: Melvin A. Breuer. Computer Sci Pr, 1975, xiii + 417 pp, $19.95. On using high level languages in the design and manufacturing of digital systems. Contents: system simulation and simulation languages, hardware design languages, their use in simulation, software design tools, and data base design. Seven co-authors. RWN

COMPUTER SCIENCE, P. *Computer Aided Design of Digital Systems, A Bibliography.* W.M. van Cleemput. Computer Sci Pr, 1976. *V. I,* xiii + 374 pp; *V. II, 1975-76,* ix + 277 pp, $30; $22 (P) each. References to books, articles, dissertations, research reports (Vol. I: through 1974; Vol. II 1975-May 1976) classified into 10 major subject headings and cross-referenced by author and key word. Annual up-dates are promised. LAS

COMPUTER SCIENCE, T(18: 1, 2), S, P. *A Theory of Programming Language Semantics.* Robert Milne, Christopher Strachey. Halsted Pr, 1976. *Part A: Indices and Appendices, Fundamental Concepts, Mathematical Foundations,* 368 pp; *Part B: Standard Semantics, Store Semantics, Stack Semantics,* 489 pp, $35. Mathematical semantics provides the approach to formalizing the semantics of programming languages. The fundamental concept in mathematical semantics is that of a valuation. The mathematical basis of the method is presented first. Elucidation is achieved through application of the method to programs written in a large and complex paradigm language. References. Appendices. Index. RJA

COMPUTER SCIENCE, T(14-17: 1), S, L. *Fundamentals of Data Structures.* Ellis Horowitz, Sartaj Sahni. Computer Sci Pr, 1976, xii + 564 pp, $16.95. Emphasizes the distinction between the specification of a data structure (expressed in formal axioms) and its realization within an available programming language. Special attention is given to designing algorithms and analyzing their computing times. Develops the ability to devise alternative implementations of a data structure. A thorough and rigorous presentation which is not intimidating. Chapter references. Exercises. Appendices. Index. RJA

COMPUTER SCIENCE, T(16-18: 1), S, L. *Computer Architecture, Second Edition.* Caxton C. Foster. Van N-Rein, 1976, xvii + 300 pp, $16.95. Expansion of the previous edition which takes recent developments in technology into account. More details on flip-flops, integrated circuits, magnetic bubble storage, semi-conductor memories, ROMS, etc. New chapter on microprocessors and up-to-date examples in the material on large computers. Examples of array processors. Exercises. Chapter references. Index. RJA

COMPUTER SCIENCE, T*(14-16: 1), S, L. *Data Structure and Management, Second Edition.* Ivan Flores. P-H, 1977, ix + 390 pp, $18.95. Straightforward, clearly organized text on data and the various ways of organizing it: files, records, fields, lists, directories, trees. Chapter on hashing and one on overflow. Many helpful examples and diagrams. Problems. Glossary. Appendix. Index. RJA

COMPUTER SCIENCE, P. *Lecture Notes in Computer Science-38: An Optimized Translation Process and Its Application to ALGOL 68.* P. Branquart, *et al.* Springer-Verlag, 1976, ix + 334 pp, $12.30 (P). Describes the translation process of a programming language into efficient machine code. The description is machine independent. Based on an actual compiler written for ALGOL 68. Bibliography. Appendices. RJA

COMPUTER SCIENCE, T*(16-17: 1), S, L. *Design of Digital Computers, An Introduction, Second Edition.* Hans W. Gschwind, Edward J. McCluskey. Springer-Verlag, 1975, viii + 548 pp, $19.50. Revision of a popular text. Three main topics: (1) number systems and Boolean algebra; (2) circuits and the basic hardware components; and (3) computer organization. Incorporates contemporary material on integrated circuits which is accessible to students with no electrical engineering or physics background. Problems. Chapter references. Index. RJA

COMPUTER SCIENCE, T(16-18: 1, 2), S, L. *Automata Theory: Machines and Languages.* Richard Y. Kain. McGraw, 1972, xvii + 301 pp, $20.50. Emphasizes the connection between machines and mathematical linguistic models. Turing machines, linear-bounded automata, pushdown automata, operations on languages, solvable and unsolvable linguistic questions. Problems. Appendices. Annotated bibliography. Index. RJA

COMPUTER SCIENCE, P. *Foundations of Computer Science II.* Ed: K.R. Apt, J.W. DeBakker. Math. Centre Tracts, No. 81 & 82. Math Centrum, 1976. *Part 1,* i + 147 pp; *Part 2,* 149 pp, Dfl. 18 (P) each. Contains the lectures of the Second Advanced Course on the Foundations of Computer Science (held May 31-June 11, 1976, at the University of Amsterdam). Part 1 lectures given by E.L. Lawler on "Graphical Algorithms and Their Complexity", and by J. VanLeeuwen on "The Complexity of Data Organization." Part 2 contains lectures by R. Milner, A. Salomaa, and W.J. Savitch. RJA

COMPUTER SCIENCE, T(16-17: 1), S, P, L. *Data Base Management Systems.* Dionysios C. Tsichritzis, Frederick H. Lochovsky. Comp. Sci. and Appl. Math. Acad Pr, 1977, xvi + 388 pp, $16.95. For the DBMS user or prospective user. Presents underlying concepts beginning with history of data processing, capabilities of DBMS, and descriptions of hierarchical, network, and relational systems. Part 2 focuses on the earlier abstractions by providing a survey of various DBMS's without dwelling on any actual current system. Exercises. Appendix. References. Index. RJA

COMPUTER SCIENCE, T(17-18: 1), S, P. *Software Portability.* Ed: P.J. Brown. Cambridge U Pr, 1977, xiv + 328 pp, $14.95. Text on the transfer of programs and data between different computers. Outgrowth of a two week intensive course at the University of Kent at Canterbury in 1976. Material organized and presented by more than a dozen experts in the field. Proceeds from introductory considerations up to consideration of two research projects. Includes legal aspects and case studies. Appendix. Index. RJA

COMPUTER SCIENCE, T(14-16: 1), S, L. *An Introduction to the Study of Programming Languages.* D.W. Barron. Cambridge U Pr, 1977, viii + 165 pp, $9.95; $4.95 (P). Considers the design principles behind the most common features found in all programming languages. Examples from a variety of popular languages illustrate each principle and provide opportunities for comparisons of the languages. Includes consideration of the interaction between language design and compiler and run-time organization. References. Exercises. Index. Index of languages. RJA

COMPUTER SCIENCE, T(13-14: 1), L. *Structured Programming and Problem-Solving with PL/1.* Richard B. Kieburtz. P-H, 1977, xiii + 348 pp, $9.95 (P). A transliteration of *Structured Programming and Problem-Solving with Algol W,* 1975 (TR, February 1976). RWN

COMPUTER SCIENCE, P. *Software Metrics.* Tom Gilb. Winthrop Pub, 1977, 282 pp, $13.95. Apparently the first book in an interesting, developing field. Part I: practical applications of software metrics. Part II: metric concepts. A general approach coupled with elaborations on the author's "Mecca" method. RWN

SYSTEMS THEORY, T(15-17: 1), S, L. *Logical Design of Digital Systems.* Arthur D. Friedman. Computer Sci Pr, 1975, x + 278 pp, $16.95. Begins with elementary material on circuits and arithmetic in various bases. Part 2 discusses combinatorial and sequential circuits in detail. Lastly, a model of a digital computer is presented, together with the accompanying problems in the physical design of a digital system. Problems. Chapter references. Appendix. Index. RJA

APPLICATIONS (ECONOMICS), P. *Lecture Notes in Economics and Mathematical Systems-141: Mathematical Economics and Game Theory, Essays in Honor of Oskar Morgenstern.* Ed: R. Henn, O. Moeschlin. Springer-Verlag, 1977, xiv + 709 pp, $22.50 (P). Papers on game theory, utility theory, economic models, economic theory, econometrics, and related topics, together with a bibliography and biography of Morgenstern. LAS

APPLICATIONS (ENGINEERING), P. *Numerical Methods in Statistical Hydrodynamics.* Alexandre Chorin. Pr U Montreal, 1977, 64 pp, $5 (P). Lecture notes. Presents the random vortex method accounting for vorticity and intermittency in two dimensions and averaging procedures which are useful for interior flows. Includes some of the preliminaries needed on probability and fluid flow. RWN

Reviewers Whose Initials Appear Above

Richard J. Allen, St. Olaf; David F. Appleyard, Carleton; Paul J. Campbell, St. Olaf; Clifton E. Corzatt, St. Olaf; John Dyer-Bennet, Carleton; Jennifer Galovich, St. Olaf; Steven Galovich, Carleton; Lorraine L. Keller, St. Olaf; Roger B. Kirchner, Carleton; Richard S. Kleber, St. Olaf; Loren C. Larson, St. Olaf; R.W. Nau, Carleton; Thomas R. Savage, St. Olaf; J. Arthur Seebach, Jr., St. Olaf; Lynn Arthur Steen, St.Olaf; T.A. Vessey, St. Olaf.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, NW, Washington, D.C. 20036. Items must be submitted at least five months before publication can take place.

PERSONAL ITEMS

Indiana University at South Bend: Associate Professor W. J. Frascella has been appointed Chairman of the Department of Mathematics; Dr. V.F. Rickey, Bowling Green State University, has been appointed Visiting Associate Professor.

Dr. G.A. Fraser, University of Santa Clara, has been promoted to Associate Professor.
Professor C. Wallace Jordan, Williams College, retired on July 1, 1977, with the title of Professor Emeritus.
Mr. Michael Kovacic, Colorado State University, has been promoted to Associate Professor.
Associate Professor D.S. Moore, Purdue University, has been promoted to Professor.

Dr. Franklyn B. Fuller, San Jose State University, died on May 12, 1977, at the age of 55. He was a member of the Association for thirty-five years.

Mr. Robert D. Lowe, Baltimore, Maryland, died on April 28, 1976, at the age of 52. He was a member of the Association for twenty-four years.

Mr. Renato L. Vitale, Riverdale, New York, died on September 5, 1976. He was a member of the Association for fifty-two years.

WHAT IS THE CONFERENCE BOARD OF THE MATHEMATICAL SCIENCES?

The Conference Board of the Mathematical Sciences (CBMS) is an organization of professional societies in the mathematical sciences serving two main purposes: (1) to provide a two-way channel of communication between the professional mathematical community represented by its member societies and the relevant Government and other organizations on the Washington scene; and (2) to function as a forum and focus for issues and projects of concern to several or all of its member societies. CBMS has two kinds of membership, constituent and affiliate. In general its constituent-member societies have primarily mathematical interests, while its affiliate-member societies have only partly mathematical interests or have grown up around special constituencies or areas of application. At present there are six constituent members: the American Mathematical Society (AMS), the Association for Symbolic Logic (ASL), the Institute of Mathematical Statistics (IMS), the Mathematical Association of America (MAA), the National Council of Teachers of Mathematics (NCTM), and the Society for Industrial and Applied Mathematics (SIAM). There are also six affiliate members: the American Statistical Association (ASA), the Association for Computing Machinery (ACM), the Association for Women in Mathematics (AWM), the Operations Research Society of America (ORSA), the Society of Actuaries (SA), and The Institute of Management Sciences (TIMS). The Conference Board has its headquarters in Washington, D.C., where it was incorporated as a non-profit educational organization in 1960. Since the fall of 1967 its headquarters office has been located at 832 Joseph Henry Building, 2100 Pennsylvania Avenue, N.W., Washington, D.C. 20037.

The CBMS role in communication between the professional mathematical community and organizations on the Washington scene is accomplished in part through direct contacts with agencies of the Federal Government and through representation on or liaison with such groups as the Committee of Scientific Society Presidents, the American Association for the Advancement of Science, the American Council on Education, the Scientific Manpower Commission, and the Office of Mathematical Sciences of the National Research Council. A principal vehicle for communication with its professional constituency is the Conference Board's *Newsletter*, published in four sixteen-page issues per year. The *Newsletter* features Washington news of interest to the broad mathematical community, notices and reports regarding national and international mathematical events, information and data on fellowships and other opportunities in mathematical research and education, manpower surveys and studies relevant for the mathematical sciences, and editorials and position papers on issues of concern to professionals in the mathematical sciences. Subscriptions to the *Newsletter* are available from CBMS at \$4.00 per year for individuals belonging to one or more member societies of CBMS and \$8.00 per year for institutions and other individuals.

The Conference Board's forum role is implemented through its semi-annual Council meetings and through the public panel discussion on some subject of interest that it regularly sponsors each year at the joint winter mathematics meeting of AMS and MAA. CBMS plans to sponsor additional such panel discussions at national meetings of its other member societies. As noted above, the editorials and position papers that appear from time to time in the CBMS *Newsletter* also contribute to this forum role. Since 1969, a major CBMS project of broad interest has been its management, under contract with the National Science Foundation, of an a· · ual program of eleven or twelve one-week Regional Conferences on subjects of current research interest in the mathematical sciences. Host institutions receiving NSF grants for such Regional Conferences in any given year are normally announced in the spring issue of the CBMS *Newsletter*. Notices of individual Regional Conferences also appear in other appropriate journals and in regional announcements by their host institutions; application for participation is to the host institution concerned. Publication, by AMS or by SIAM, of monographs resulting from the Regional Conferences is arranged by CBMS. Other CBMS projects of broad interest have included a major study of information-service needs of the mathematical sciences; a series of surveys of undergraduate and graduate mathematical education and a survey of school-level mathematical education; an earlier study of buildings and facilities for the mathematical sciences; and projects on public understanding of the mathematical sciences and their applications.

The historical antecedents of the Conference Board go back to 1942, when a War Policy Committee was formed by AMS and MAA with Rockefeller Foundation funds. This continued following World War II as a Policy Committee for Mathematics which by 1957 included all the present constituent members of CBMS. In 1958 it became the Conference Organization of the Mathematical Sciences with a formal constitution and by-laws and with a Washington headquarters office established in 1959 through a grant from the Carnegie Corporation to MAA. It was after its formal incorporation in 1960 in the District of Columbia as the Conference Board of the Mathematical Sciences that its present affiliate members joined: ACM in 1962; ORSA, SA and TIMS in 1966; ASA in 1973; and AWM in 1976.

TRUMAN BOTTS, *Executive Director, CBMS*

SMITHSONIAN INSTITUTION FELLOWSHIPS

The Smithsonian Institution will award fellowships for research in residence at the I··stitution during 1978–79 in a number of scientific fields, one of which is History of Mathematics. For information write to Gretchen Gayle, Program Officer, Office of Academic Studies, Smithsonian Institution, Washington, D. C. 20560.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

APRIL MEETING OF THE ALLEGHENY MOUNTAIN SECTION

The annual spring meeting of the Allegheny Mountain Section of the MAA was held at St. Francis College, Loretto, PA, April 22 and 23, 1977. The host was Professor Adrian Baylock of St. Francis College. There were approximately 75 members of the Association in attendance along with 50 students of the member schools.

The program was divided into two parts, one faculty and one student. The major faculty address Friday evening was presented by Julia Knight, Pennsylvania State University, entitled "Models of Arithmetic". This was followed by a panel on "Remedial Mathematics" moderated by Earle Myers, University of Pittsburgh. The panel members included Mary Bivens, Allegheny College; Frank Hiergeist, West Virginia University; and Kay Hudspeth, Pennsylvania State University. The keynote student address was given by Barbara Faires, Westminster College, entitled "Mathematical Puzzles". Her students at Westminster assisted in the presentation.

The following contributed talks were presented:

FACULTY

1. *Integral representations of* $SL(2, C)$, by E.R. Anderson, West Virginia Wesleyan College.
2. *A schedule for a double elimination tournament*, by E.W. Klaber, Washington & Jefferson College.
3. *Permutation polynomials in several variables over a finite field*, by Gary Mullen, Pennsylvania State University, Shenango Valley Campus.

4. *A note on odd abundant numbers*, by C.A. Cable, Allegheny College.
5. *Lucas triangle primality criterion dual to that of Mann–Shanks*, by H. W. Gould and W.E. Greig, West Virginia University.
6. *Mathematics for health sciences*, by Kalman Mecs, Allegheny Campus, Community College of Allegheny County.
7. *Calculus for students who need HELP in algebra*, by F.H.S. Hall, Pennsylvania State University, Fayette Campus.
8. *A trigonometric treatment of a maximum*, by J.C. Eaves, West Virginia University.
9. *Summer program for high-risk college students at Edinboro State College*, by J.D. Urban, Edinboro State College.

STUDENT

1. *An application of the Burnside counting theorem*, by James Bretti, Allegheny College.
2. *The groups of collineations of a nine-point geometry*, by Michael Monnett, Allegheny College.
3. *Catalan structures and correspondences*, by Michael Kuchinski, West Virginia University.
4. *The W-function: An experiment in elementary number theory*, by Edward Weismann, Allegheny College.
5. *Constructing norms in vector spaces*, by Robert Goldrick, Allegheny College.
6. *Regular languages and automata*, by Suzanne Rex, West Virginia Wesleyan College.
7. *Z-parameters*, by James Snyder, Butler County Community College.
8. *Derivation of the class equation*, by Clayton Tenney, West Virginia Wesleyan College.
9. *Maximum power transfer theory*, by Douglas Bartley, Butler County Community College.
10. 0^0 - *The puzzler*, by Brother Joe Robinson, T.O.R., St. Francis Seminary.
11. *Mathematical terminology—A humorous view*, by David Loth, St. Francis College.

The Saturday morning session was devoted to three invited addresses. These were:

1. *The history of the four-color problem*, by Eric Braude, Pennsylvania State University, Behrend College.
2. *Is addition really commutative?*, by Richard McDermot, Allegheny College.
3. *Prime generating functions and congruences*, by Henry Alder, University of California, Davis, and President of the MAA.

Professor James Derr, Chairman of the Section, presided at the business meeting. The secretary's report included a list of the four students in the Allegheny Mountain Section of the MAA. The top four students were J.E. Brosius, Penn State; Robert Caflisch, West Virginia University; David Housman, Allegheny College; and Y.B. Cohen, Carnegie-Mellon University. A one-year subscription to *Mathematics Magazine* will be awarded to these students. A report on the High School Mathematics Contest was given by Professor I.D. Peters, West Virginia University, for West Virginia, and by Professor F.H.S. Hall, Penn State University, Fayette Campus, for Western Pennsylvania. The sectional governor, Earle Myers, University of Pittsburgh, reported on the Speakers Bureau and invited the section to the University of Pittsburgh for next year's meeting. A report was given on the state of the MAA by the President of the MAA, Henry Alder.

Professor Derr made an announcement concerning the short course on Mathematical Modeling to be held at Ohio University, June 14-18, 1977, which the Allegheny Mountain Section is co-sponsoring with the Ohio Section. The nominating committee for 1977-78 was named to be I.D. Peters, West Virginia University; James Reynolds, Penn State, Beaver Campus; Robert Anderson, West Virginia Wesleyan College; and Kathleen Taylor, Duquesne University.

Officers elected were Richard McDermot, Allegheny College: Chairman; Carol Booth, West Liberty State College: Second Vice-Chairman; and Richard Lundgren, Allegheny College: Coordinator of the Student Program. Frank Kocher, Pennsylvania State University, will serve as First Vice-Chairman. Earle Myers, University of Pittsburgh, will continue to serve as Governor of the Allegheny Mountain Section, and John Milsom, Butler County Community College, will serve as Secretary-Treasurer.

J.W. MILSOM, *Secretary-Treasurer*

APRIL MEETING OF THE IOWA SECTION

The 64th regular meeting of the Iowa Section of the MAA was held on the campus of Drake University, Des Moines, on April 23, 1977. Chairman James Cornette presided. There were 49 in attendance, 39 of whom were members of the section. For the invited address, sponsored jointly by MAA and the Iowa Academy of Sciences, there were 60 persons present.

The program consisted of the contributed papers, an invited address by Professor R.P. Boas, Governor's report by Elsie Muller, and the business session. D.V. Meyer, Central College, Pella, was elected as the Chairman-elect.

The program, arranged by Ellen Oliver, consisted of the following:

1. *Reciprocals of integers as repeating decimals*, by B.E. Gillam, Des Moines.
2. *Modular instruction in Mathematics at Central College*, by Leland Graber, Pella.
3. *The right constructs for programming*, by R.F. Keller, Ames.
4. *Staying alive as a mathematician while teaching at a two-year college*, by R.H. Lambertson, Boone.
5. *The programmable calculator as a teaching device*, by Don Benbow, Norlin Rober, Don McVay, Marshalltown.
6. *A mathematical analysis of the metropolitan government in Polk County, Iowa*, by A.F. Kleiner, Jr., Des Moines.
7. *Positive invariant closed sets for delay differential equations*, by Y.F. Chang, Ames.
8. *Find polynomial inverses of matrices of rational functions*, by James Bruening, LeMars.

Invited Address: *The Harmonic series and some of its applications*, by R.P. Boas, Past-President, MAA, Editor, American Mathematical Monthly, Northwestern University, Evanston, Illinois.

B.E. GILLAM, *Secretary-Treasurer*

APRIL MEETING OF THE ROCKY MOUNTAIN SECTION

The Sixtieth Annual Meeting of the Rocky Mountain Section of the MAA was held on the campus of Metropolitan State College, Denver, Colorado, April 29-30, 1977. There were 143 registrations, including Forest Fisch, Governor of the Section, Donald Bushnell, Chairman of the Section, and H.O. Pollak of Bell Laboratories. The Banquet address "On the Relationship between Applications of Mathematics and the Teaching of Mathematics" was delivered by Dr. Pollak. Invited addresses were: "A Problem-Solving Course in an Industrial Setting", by Armando Gingras, Metropolitan State College; "A Transition Course Between Calculus and Upper Division Mathematics", by Kent Goodrich, University of Colorado; and "On Shortest Connecting Networks", by H.O. Pollak.

There were twenty-six contributed papers:

1. *Those ubiquitous Catalan numbers*, by R.A. Gibbs, Fort Lewis College.
2. *On the conjugates of* $1-2|x|$, by Jan Mycielski, University of Colorado.
3. *Puzzles and graphs*, by Stan Gudder, Denver University.
4. *A radical theory for hemirings and an open problem*, by D.M. Olson, Cameron University.
5. *Certain matrix congruences and* $(\mathrm{mod}\, p^n)$, by T.P. Donovan, University of Colorado.
6. *The conjecture* $\pi(x+y)<\pi(x)+\pi(y)$, by David Ballew, South Dakota School of Mines and Technology.
7. *Ranked solutions of some matrix equations*, by Nick Mousouris, Humboldt State University.
8. *An extremum problem in two independent variables*, by R.S. Fisk, Colorado School of Mines.
9. *Some forgotten topics in analytic geometry*, by D.W. Hardy, Colorado State University.
10. *Sex differences in the learning of mathematics*, by Nancy Angle, University of Colorado at Denver.
11. *On residue class cryptography*, by Ron Whittekin, Metropolitan State College.
12. *Another construction of a strictly increasing, continuous, singular function on* $[0,1]$, by Paul O'Meara, University of Colorado at Denver.
13. *Computer arithmetic and abstract algebra*, by Keith Joseph, Metropolitan State College.
14. *Problem-solving techniques usable in secondary mathematics*, by Melfried Olson, University of Wyoming.
15. *An assessment of the processes used by community college students in mathematical problem-solving with suggestions for teaching*, by Beverly Gimmestad, Metropolitan State College.
16. *Emerging trends in secondary teacher education*, by Earl Hasz, Metropolitan State College.
17. *Some simple applications of statistics in industry*, by M.F. Flynn, Coors Container Company.
18. *Research suggestions in Lanchester combat theory*, by P.M. Ellis, Utah State University.
19. *Mathematical programming solutions to identification problems in mass spectroscopy*, by D.W. Fausett, Colorado School of Mines.

20. *On infinite games, the law of large numbers, and Baire category*, by Celestino Mendez, Metropolitan State College.

21. *A cooperative audio-visual project in mathematics and electrical engineering*, by David Ballew and Ronald Schmitz, South Dakota School of Mines and Technology.

22. *Research results with implications for the teaching of general math*, by Bill Juraschek, University of Colorado at Denver.

23. *Matrices can be useful (and fun)!!!*, by A.D. Porter, University of Wyoming.

24. *The curve parallel to a parabola is not a parabola*, by F.M. Stein, Colorado State University.

25. *Problems in teaching the history of mathematics*, by Burnett Meyer, University of Colorado.

26. *Bye-bye slide rule*, by L.M. Orman, University of Southern Colorado.

In addition to the above papers and addresses, exhibits were presented by the Association for Women in Mathematics, McGraw-Hill Book Company, Macmillan Publishing Company, Inc., Prindle, Weber & Schmidt, Inc., MAA, and Worth Publishers, Inc.

The business meeting was convened on Saturday morning, April 30, 1977, at 8:00 A.M. by Professor Bushnell who presided. Thirty-eight members attended. The minutes of the 1976 meeting were circulated and approved. The Treasurer's report for 1975 and an interim report for 1976 were circulated and approved.

Professor Dean Benson presented the Nominating Committee's slate of nominees; the following officers were elected: Vern Nelson, Metropolitan State College, Chairman; John Hodges, University of Colorado, Chairman-Elect; C.A. Grimm and Dale Rognlie, South Dakota School of Mines and Technology, Program Chairmen.

Christopher Bretherton, of the University of Colorado, and Peter Li, of the University of Southern Colorado, were given memberships to the MAA for their high placement on the Putnam Examination.

Professor Forest Fisch, University of Northern Colorado, gave the Governor's report from the Toronto and St. Louis meetings. Professor George Donovan, Metropolitan State College, Chairman of the High School Lectureship program, presented that committee's report. The Section contributed $100 to the continuance of that program. A written report from the High School Mathematics Contest had been submitted by Professor Robert Vunovich of the University of Southern Colorado.

Professor Vunovich was commended for his service to the Mathematics Contest.

The Section voted to start a newsletter edited by Professor David Ballew, South Dakota School of Mines and Technology.

The meeting concluded with a "rap session" conducted by Dr. H.O. Pollak, Past President of the MAA.

DAVID BALLEW, *Secretary-Treasurer*

APRIL MEETING OF THE WISCONSIN SECTION

The 1977 annual meeting of the Wisconsin Section of the MAA was held on the campus of the University of Wisconsin at Oshkosh on April 29 and 30, 1977. There were 135 registered attendants, including 112 MAA members and 23 students.

At the business meeting some revisions in the section by-laws were passed. The section voted to begin a yearly section newsletter. Based on the success of our first fall workshop on mathematical models, the Section voted to hold a second fall workshop.

The two principal speakers at the meeting were Professor Wolfgang Haken of the University of Illinois and Professor Leonard Gillman of the University of Texas. Professor Haken's talk was: "The Four-Color Problem"; Professor Gillman's talk was: "Choosing a Wife."

The contributed presentations were:

The human side of Gauss, by Merrill Barnaby, University of Wisconsin-LaCrosse.

Sylvester at Hopkins: an American centennial, by John Finch, Beloit College.

The misnamed Gaussian (normal) curve, by Carroll Rusch, University of Wisconsin-Superior.

The pocket calculator in higher education, by Eli Maor, University of Wisconsin-Eau Claire.

Returning theory of equations to the undergraduate curriculum, by C.W. Schelin, University of Wisconsin-LaCrosse.

Trees and maximal outplanar graphs, by A.E. Barkauskas, University of Wisconsin-LaCrosse.

Zeeman's catastrophe machine, by Wilbur Hoppe, University of Wisconsin-Eau Claire.

Ice crystal halos, by Walt Tape, University of Wisconsin-Eau Claire.

Special session on the problem of poor math preparation of students entering college, by J.C. Neuenfeldt, University of Wisconsin-Stout; Sam Filippone, University of Wisconsin-Parkside; and Gary Klatt, University of Wisconsin-Whitewater.

Derivation of Simpson's rule from a comparison of the midpoint and trapezoidal rules, by Walt Sadler, Universisty of Wisconsin-Waukesha.

Can computers illustrate calculus?, by J.H. Tutsch, University of Wisconsin-Marathon County.

The use of computers in the teaching of introductory statistics, by R.P. Situmeang, University of Wisconsin-LaCrosse.

An exercise approach to calculus with the computer, by Tim Fossum and Ron Gatterdam, University of Wisconsin-Parkside.

Minitab—a simple statistical computing system, by R.L. Andrews and Ann Goodsell, University of Wisconsin-Oshkosh.

Classical boundary value problems and computer graphics: a class project in applied mathemtics, by Eric Ahlvin, Larry Bruno, Dennis Grau, Luther Johnson, Marge Stankus, Steve Weingarth, Michael Welcome, (students) and Donald Piele (instructor), University of Wisconsin-Parkside.

TOM RENFROW, *Secretary-Treasurer*

1977 CONTRIBUTING MEMBERS AND SPECIAL GIFTS

The Association is deeply indebted to the generosity of the 119 members listed below who have elected to be Contributing Members, Sponsors, or Patrons for 1977 by making contributions beyond the normal dues.

Patrons: Leonard Gillman, Peter M. Goetschius, Herbert J. Ryser, Saturnino L. Salas, Earl W. Swokowski.

Sponsors: William W. Babcock, George H. Bridgman, William G. Chinn, Richard A. Cleveland, H. J. Ellsworth, Andrew M. Gleason, William E. Hartnett, W. R. Harris, Bill Hassinger, Jr., Paul B. Johnson, Warren S. Loud, Gene M. Ortner, H. O. Pollak, Maria W. Steinberg, Robert C. Wever.

Contributing Members: H. L. Alder, E. Z. Andalafte, R. D. Anderson, T. M. Apostol, J. A. Ball, J. R. Baugh, E. M. Beesley, L. Bernstein, W. M. Boyce, M. K. Brachman, M. H. Broberg, Mildred C. Brunschwig, G. C. Byers, F. P. Cass, P. M. Chase, A. H. Clifford, H. D. Colson, L. C. Corey, Helen F. Cullen, A. B. Cunningham, H. F. Davis, C. R. Deeter, W. E. Deskins, W. J. Dodge, R. Dubisch, H. V. Ellis, E. O. Embree, R. J. Engert, T. Evans, I. M. Ferris, A. M. Fischer, H. S. Fox, R. E. Gaskell, J. H. Gehrke, Elaine F. Genkins, M. J. Gilpin, M. A. Glass, R. Goldberg, R. A. Good, W. H. Gottschalk, J. H. Graham, D. A. Gregory, W. L. Hart, J. Hashisaki, P. Henrici, P. S. Herwitz, F. E. Hohn, J. M. Horvath, M. Gweneth Humphreys, T. E. Ikard, C. A. Jacewicz, S. B. Jackson, M. P. Jarnagin, J. L. Johnson, W. Kaplan, J. E. Kiefer, D. E. Knuth, R. M. Krause, D. L. Kreider, H. A. Krieger, R. P. Kroeckel, J. C. Lamerand, L. H. Lange, L. H. Lanier, Jr., J. P. LaSalle, Diana Lieu, J. A. Long, Barbara Louthan, W. R. Mann, M. J. Mansfield, W. Margulies, J. R. Mayor, R. K. McConnell, Jr., R. M. McLeod, P. L. Montgomery, R. A. Moore, S. A. Naimpally, D. A. Nelson, C. J. Nesbitt, A. Nijenhuis, E. J. Pellicciaro, J. Pitcairn, M. H. Protter, M. O. Reade, C. E. Richart, D. C. Rose, D. P. Roselle, J. R. Schap, N. E. Sexauer, A. L. Shields, D. B. Singmaster, R. H. Sorgenfrey, J. B. Starkey, J. H. Taub, C. R. Tellefsen, S. Wallach, A. J. Washington, F. J. Weyl, R. E. Zink.

The Association also acknowledges with thanks the following special gifts:

A bequest of $47,462 from the estate of Carl B. Allendoerfer, 28th President of the Association.

A gift of $2,538 from Dorothy Allendoerfer.

A gift of $50,000 from Carol Ryrie Brink in memory of Raymond W. Brink, 19th President of the Association.

A gift of $70,000 from Mary P. Dolciani to establish a Building Fund for a future MAA Headquarters Building.

Anonymous contributions of $500 and $25 from members. A gift of $25 from a non-member.

CALENDAR OF FUTURE MEETINGS

Sixty-first Annual Meeting, Atlanta, Georgia, January 6-8, 1978.

Fifty-eighth Summer Meeting, Brown University, August 8-10, 1978.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN, University of Pittsburgh, Pennsylvania, April 14–15, 1978.

FLORIDA, St. Petersburg Junior College, Clearwater, March 3-4, 1978.

ILLINOIS, Western Illinois University, Macomb, May 5-6, 1978.

INDIANA, Indiana Central College, Indianapolis, November 5, 1977.

INTERMOUNTAIN

IOWA, University of Northern Iowa, Iowa Falls, April 22, 1978.

KANSAS, Wichita State University, Wichita, late March–early April 1978.

KENTUCKY, early April. Deadline for papers 6 wks. bef. mtg.

LOUISIANA-MISSISSIPPI, Buena Vista Hotel-Motel, Biloxi, Mississippi, February 17-18, 1978.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA, American University, Washington, D.C., November 19, 1977.

METROPOLITAN NEW YORK, late April or early May 1978. Deadline for papers 2 wks. bef. mtg.

MICHIGAN, Michigan State University, East Lansing, Spring 1978.

MISSOURI, Central Missouri State University, Warrensburg, April 7-8, 1978.

NEBRASKA, University of Nebraska at Omaha, April 14-15, 1978.

NEW JERSEY, Caldwell College, Caldwell, November 5, 1977.

NORTH CENTRAL, College of St. Thomas, St. Paul, Minnesota, April 21–22, 1978.

NORTHEASTERN, Saturday after Thanksgiving, and third week in June in odd-numbered years.

NORTHERN CALIFORNIA, College of Notre Dame, Belmont, February 18, 1978.

OHIO

OKLAHOMA-ARKANSAS, Henderson State University, Arkadelphia, Arkansas, March 31-April 1, 1978.

PACIFIC NORTHWEST, University of Oregon, Eugene, June 16–17, 1978.

PHILADELPHIA, Moravian College, Bethlehem, November 19, 1977.

ROCKY MOUNTAIN, South Dakota School of Mines and Technology, Rapid City, April 28-29, 1978.

SEAWAY, first Saturday in November and Saturday in late April. Deadline for papers 6 wks. bef. mtg.

SOUTHEASTERN, Clemson University, Clemson, South Carolina, March 31-April 1, 1978.

SOUTHERN CALIFORNIA, California State Polytechnic University, San Luis Obispo, November 11-12, 1977.

SOUTHWESTERN, New Mexico Institute of Mining and Technology, Socorro, Spring 1978.

TEXAS, Stephen F. Austin State University, Nacogdoches, March 31-April 1, 1978.

WISCONSIN, University of Wisconsin-Whitewater, late April 1978.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Washington, February 12-17, 1978.

AMERICAN MATHEMATICAL ASSOCIATION OF TWO YEAR COLLEGES

AMERICAN MATHEMATICAL SOCIETY, Atlanta, Georgia, January 4-7, 1978.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION

ASSOCIATION FOR COMPUTING MACHINERY

ASSOCIATION FOR SYMBOLIC LOGIC

ASSOCIATION FOR WOMEN IN MATHEMATICS, Atlanta, Georgia, January 4–8, 1978.

CANADIAN SOCIETY FOR HISTORY AND PHILOSOPHY OF MATHEMATICS/SOCIÉTÉ CANADIENNE D'HISTOIRE ET DE PHILOSOPHIE DES MATHÉMATIQUES

FIBONACCI ASSOCIATION

INSTITUTE OF MATHEMATICAL STATISTICS

MU ALPHA THETA

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, San Diego, California, April 12-15, 1978.

OPERATIONS RESEARCH SOCIETY OF AMERICA, Peachtree Plaza Hotel, Atlanta, November 7-9, 1977.

PI MU EPSILON

SCHOOL SCIENCE AND MATHEMATICS ASSOCIATION, William Penn Hotel, Pittsburgh, November 10-12, 1977.

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, Hilton Inn, Albuquerque, New Mexico, October 31-November 2, 1977.

The best math texts are those that WORK!

And these four from P-H have *proven* classroom performance.

CALCULUS WITH ANALYTIC GEOMETRY, 3rd Edition

Edwin J. Purcell — Emeritus, University of Arizona

Third Edition of highly successful introductory text is now better than ever. Bright new design includes open, attractive pages and many new detailed, illustrative examples. Written with the average student in mind. Unnecessary verbiage is avoided and the explanations are clear, straight-forward, and simple. Features more motivational material introducing concepts and better transitional material between chapters than previous editions. Exercises have been revised and reorganized to proceed from the easy to the more difficult. Students can read and master Purcell alone, or with minimal instruction. Also new to this edition is the Solutions Manual — available to instructors and students, if desired. For a two- or three-semester course.

1978 **960 pp. (est.)** **Cloth $19.95**

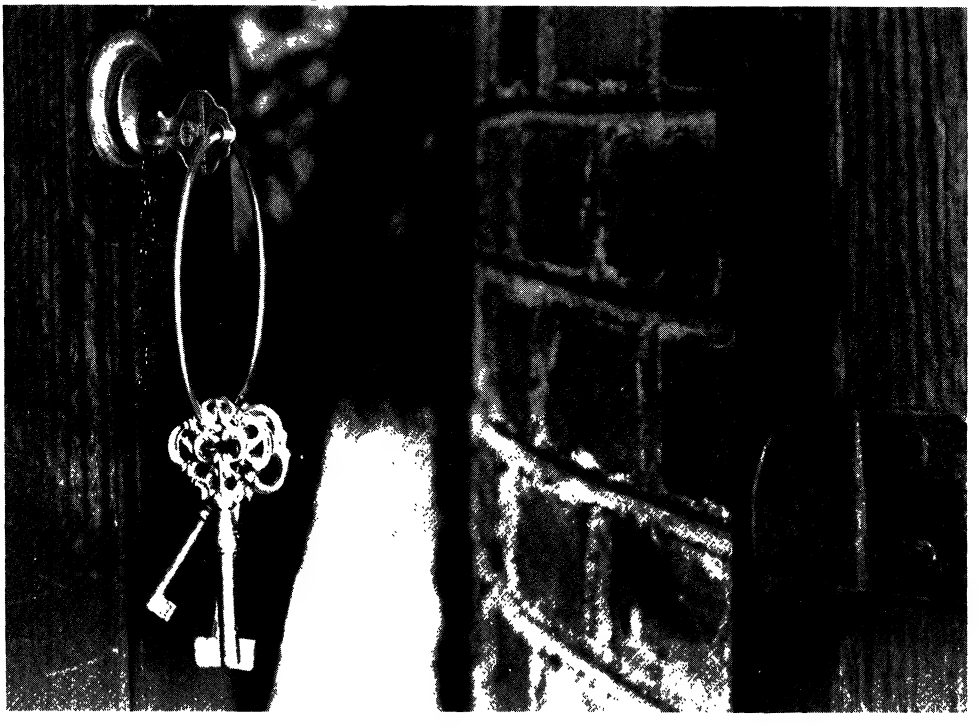
INTRODUCTION TO MATHEMATICS, 4th Edition

Bruce E. Meserve — University of Vermont;
Max A. Sobel — Montclair State University

Survey of contemporary mathematics designed for liberal arts, education, and business students presents new topics and explores varied ways of looking at mathematics. Fourth Edition features a new attractive and motivational design. New chapters on the metric system and computers have been added, and the chapter on geometry has been revised and oriented toward liberal arts students. "Mathematical Explorations" and "Readings and Projects" sections provide supplementary material for class discussion and individual projects.

1978 **512 pp. (est.)** **Cloth $12.95**

Keyed to fit



your curriculum

GOODMAN: Calculus for the Social Sciences

Now you can give your non-mathematics majors the key to basic calculus without overburdening them with proving theorems. Calculus for the Social Sciences is an innovative approach to the subject, that avoids proofs of obvious results (limiting some to the appendix or others to exercises.)

Throughout, emphasis is on practicality and realistic examples. A two-color format highlights key formulas and equations and most chapters include review questions and problems. Answers to all questions and problems are listed in the back of the book.

The text is readily adaptable to either a one- or two-semester course. The more difficult material is reserved for appendices on inequalities, theorems, implicit functions and differentiation, the height of a curve, and the area of a surface. Tables of squares and square roots, logarithms, exponential functions and integrals are also included.

By **A.W. Goodman**, University of South Florida. 442 pp. 118 ill. $12.95. Jan. 1977.

BELLO: Algebra for College Students

Here's a text that can be used independently in both community and four-year college intermediate programs. Its systematic progression of topics will lend itself easily to your individual pace of instruction. Plus, students will appreciate the use of a second color in margin notes, rules, and definitions. A Student Study Guide is also available.

By **Ignacio Bello,** Hillsborough Community College, Tampa. 701 pp. 101 ill. $14.95. May 1977.

SETEK: Algebra: A Fundamental Approach

The author supplies extensive explanations of key concepts supported by numerous worked examples graded according to difficulty. Students will then complete their mastery of the topic through completion of the 40-60 problems supplied with each section. It's perfect for the four-year college or community college student.

By **William M. Setek, Jr.,** Monroe Community College, Rochester. 708 pp. Illustd. $12.95. March 1977.

BENNETT, MILLER & STEIN: Plane Trigonometry

Heavily illustrated throughout with diagrams, charts, and boxed theorems and definitions, this text provides a thoroughly accessible introduction to college trigonometry with plenty of sophisticated, real world applications. Each chapter contains review problems. Practice exams and summation of results are also included.

By **Michael E. Bennett, Richard A. Miller** and **Barry N. Stein,** all of the Bronx Community College of CUNY. 430 pp. 316 ill. $11.50. March 1977.

SILVERMAN: Essential Calculus with Applications

This short approach to calculus supplies all the basics that can be mastered in a one-semester course. Here the key is modern applications of calculus with more than 100 pages of real life problems. A Teacher's Guide is free upon adoption.

By **Richard A. Silverman,** New School for Social Research. 267 pp. Illustd. $10.95. Feb. 1977.

PAYNE: Pre-Calculus Mathematics

Ideal for students at either community or four-year colleges, this text can also be used in the algebra-trigonometry sequence that leads to calculus. Whatever the course, you'll find that your students will benefit from the clear, extensive explanations of concepts. An Instructor's Guide is free upon adoption.

By **Michael Payne,** PhD, College of Alameda. 429 pp. 210 ill. $12.95. April 1977.

SELLERS: Elementary Statistics

Introduce your students to statistics with this attention-getter. A one-semester text, it covers all the basic topics (including probability and non-parametric statistics), while emphasizing descriptive statistics. Also, you'll find many charts, diagrams, and newspaper clippings illustrating key points. Plus, there are numerous real-world exercises and problems.

By **Gene R. Sellers,** Sacramento City College. 433 pp. 364 ill. $12.50. April 1977.

NOSAL: Basic Probability and Applications

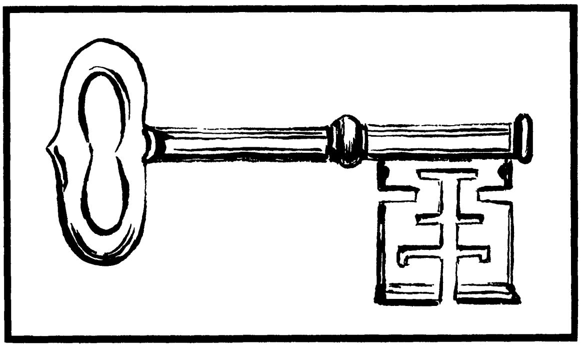
Now you can give your students a sound introduction to probability without assuming a calculus background. Nosal's comprehensive text works for non-majors and majors alike because of its logical progression of topics.

By **Miloslav Nosal,** Associate Professor, Dept. of Mathematics and Statistics, Univ. of Calgary. 370 pp. $13.50. June 1977.

COOPER, BHAT & Le BLANC: Introduction to Operations Research Models

The complex stragegies of model building in operations research, applied mathematics and management science are brought to terms your students will readily understand in this effective study. Actual problem solving situations trace the development of each model and show its applicability in business and industry.

By **Leon Cooper, U. Narayan Bhat,** and **Larry J. Le Blanc,** all of the School of Engineering and Applied Science, Southern Methodist Univ. 404 pp. Illustd. $14.95. March 1977.



W.B.Saunders Company

W. Washington Square
Philadelphia, Pa. 19105

Just published—the new

INDEX OF THE AMERICAN MATHEMATICAL MONTHLY

Volumes 1 through 80 (1894–1973)

Edited by KENNETH O. MAY

This is a conventional index to articles and a selection of other items in the first eighty volumes of the American Mathematical Monthly (1894–1973). It consists of a chronological table of contents of articles indexed, an author index to these articles, and a subject index to them and a few items not listed in the Contents. The indexes are like those at the backs of books and should be used in the same way.

Table of Contents, Author Index, Subject Index: vi + 269 pages.

Individual members of the Association may purchase one copy of the book for $10.00; additional copies and copies for nonmembers are priced at $16.00 each. (Orders for under $10.00 must be accompanied by payment. Prepaid orders will be delivered postage and handling free.)

Orders should be sent to:

MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.
Washington, D.C. 20036

U.S. POSTAL SERVICE

STATEMENT OF OWNERSHIP, MANAGEMENT AND CIRCULATION

(Required by 39 U.S.C 3685)

1 TITLE OF PUBLICATION: AMERICAN MATHEMATICAL MONTHLY

A PUBLICATION NO: A M M Y A E

2 DATE OF FILING: Sept. 1, 1977

3 FREQUENCY OF ISSUE: monthly except July, August

A NO. OF ISSUES PUBLISHED ANNUALLY: ten

B ANNUAL SUBSCRIPTION PRICE: $25 00

4 LOCATION OF KNOWN OFFICE OF PUBLICATION *(Street, City County State and ZIP Code) (Not printers)*: 1225 Connecticut Ave. NW, Washington, D.C. 20036

5 LOCATION OF THE HEADQUARTERS OR GENERAL BUSINESS OFFICES OF THE PUBLISHERS *(Not printers)*: 1225 Connecticut Ave. NW, Washington, D.C. 20036

6 NAMES AND COMPLETE ADDRESSES OF PUBLISHER, EDITOR AND MANAGING EDITOR

PUBLISHER *(Name and Address)*: The Mathematical Association of America, 1225 Connecticut Ave. NW, Washington, DC 20036

EDITOR *(Name and Address)*: Alex Rosenberg, Cornell Univ., Ithaca, NY and R. P. Boas, Northwestern Univ.,Evanston, IL

MANAGING EDITOR *(Name and Address)*: Dr. Raoul Hailpern, Math. Ass'n. of America, SUNY at Buffalo, Buffalo, NY 14214

7 OWNER *(If owned by a corporation, its name and address must be stated and also immediately thereunder the names and addresses of stock holders owning or holding 1 percent or more of total amount of stock. If not owned by a corporation, the names and addresses of the individual owners must be given. If owned by a partnership or other unincorporated firm, its name and address, as well as that of each individual must be given.)*

NAME	ADDRESS
The Mathematical Association of America, Inc.	1225 Connecticut Ave. NW Washington, DC 20036

8. KNOWN BONDHOLDERS, MORTGAGEES, AND OTHER SECURITY HOLDERS OWNING OR HOLDING 1 PERCENT OR MORE OF TOTAL AMOUNT OF BONDS, MORTGAGES OR OTHER SECURITIES *(If there are none, so state)*

NAME	ADDRESS
none	

9 FOR COMPLETION BY NONPROFIT ORGANIZATIONS AUTHORIZED TO MAIL AT SPECIAL RATES *(Section 132.122 PSM)*
The purpose, function, and nonprofit status of this organization and the exempt status for Federal income tax purposes *(Check one)*

[X] HAVE NOT CHANGED DURING PRECEDING 12 MONTHS [ ] HAVE CHANGED DURING PRECEDING 12 MONTHS *(If changed, publisher must submit explanation of change with this statement.)*

10 EXTENT AND NATURE OF CIRCULATION	AVERAGE NO. COPIES EACH ISSUE DURING PRECEDING 12 MONTHS	ACTUAL NO. COPIES OF SINGLE ISSUE PUBLISHED NEAREST TO FILING DATE
A TOTAL NO. COPIES PRINTED *(Net Press Run)*	21,750	21,000
B PAID CIRCULATION 1. SALES THROUGH DEALERS AND CARRIERS, STREET VENDORS AND COUNTER SALES	none	none
2. MAIL SUBSCRIPTIONS	19,204	18,587
C TOTAL PAID CIRCULATION *(Sum of 10B1 and 10B2)*	19,204	18,587
D FREE DISTRIBUTION BY MAIL, CARRIER OR OTHER MEANS SAMPLES, COMPLIMENTARY, AND OTHER FREE COPIES	35	7
E TOTAL DISTRIBUTION *(Sum of C and D)*	19,239	18,594
F COPIES NOT DISTRIBUTED 1. OFFICE USE, LEFT OVER, UNACCOUNTED, SPOILED AFTER PRINTING	2,511	2,406
2. RETURNS FROM NEWS AGENTS	none	none
G TOTAL *(Sum of E, F1 and 2—should equal net press run shown in A)*	21,750	21,000

11 I certify that the statements made by me above are correct and complete

SIGNATURE AND TITLE OF EDITOR, PUBLISHER, BUSINESS MANAGER, OR OWNER: Executive Director

12. FOR COMPLETION BY PUBLISHERS MAILING AT THE REGULAR RATES *(Section 132.121, Postal Service Manual)*

39 U. S. C. 3626 provides in pertinent part: "No person who would have been entitled to mail matter under former section 4359 of this title shall mail such matter at the rates provided under this subsection unless he files annually with the Postal Service a written request for permission to mail matter at such rates."

In accordance with the provisions of this statute, I hereby request permission to mail the publication named in Item 1 at the phased postage rates presently authorized by 39 U. S. C. 3626

SIGNATURE AND TITLE OF EDITOR, PUBLISHER, BUSINESS MANAGER, OR OWNER

PS Form Mar 1977 3526 *(Page 1)* (See instructions on reverse)

Sixth Edition 1975

GUIDEBOOK

TO

DEPARTMENTS IN THE MATHEMATICAL SCIENCES IN THE UNITED STATES AND CANADA

. . . intended to provide in summary form information about the location, size, staff, library facilities, course offerings, and special features of both undergraduate and graduate departments in the Mathematical Sciences . . .

100 pages, 1350 entries.

Price: $3.00

Orders with remittance should be sent to:

MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, NW
Washington, D.C. 20036

FIRST ANNUAL WORKSHOP ON
THE INFORMATION LINKAGE BETWEEN APPLIED MATHEMATICS AND INDUSTRY

to be held at

Naval Postgraduate School, Monterey, California
February 23 through 25

PRELIMINARY ANNOUNCEMENT AND CALL FOR PAPERS

It has been observed that a number of projects in government laboratories and industry could benefit from greater application of the latest developments in the field of mathematics. Secondly, some observers feel that "applied" mathematics is not always developing techniques that apply to current problems.

The objectives of this workshop are: (a) to direct the attention of applied mathematicians to some of the real problems that engineers in government laboratories and industry face and need solved, (b) to acquaint engineers in laboratories and industry with recent advances in mathematics that have been effective and are promising.

The workshop will be divided into two parts consisting, respectively, of invited papers and contributed papers.

Part I—Invited Papers: Sessions in each of the following disciplines:

I. Large Scale Computers and Traditional Applied Mathematics
II. Optimization
III. Stochastic Modeling
IV. Finite Mathematics

In each session two outstanding workers—an applied mathematician and an engineer—will start off the session by speaking jointly on a common topic of general interest within the concerned discipline. The engineer will present his problems, techniques, constraints, difficulties and trends, while the mathematician will present recent developments that are useful or show promise. This initial topic will lead to other areas of interest within the discipline. The ensuing discussions by the principal speakers, as well as by participants from the audience should be of interest to everyone. Engineers will be introduced to recent mathematical developments and meet mathematicians with ideas and interest in modeling and solving their problems. Mathematicians also will find problems which may require new techniques and methods and establish contacts with professional colleagues who understand these problems in their larger setting.

Included among the invited speakers are: Herbert Keller (California Institute of Technology), Carol Newton (UCLA), Victor Klee (University of Washington), James McKenna (Bell Laboratories), Larry Bruckner (ERDA), Alan Konheim (IBM), and Herman Chernoff (MIT).

Part II—Contributed Papers: Authors are invited to submit papers in engineering or applied mathematics with abstracts of not more than one page including title. The closing date for abstracts is January 1, 1978. Notice of acceptance will be mailed about January 15, 1978.

Send two copies of abstracts, typed in double space and inquiries for further information to:

Workshop Committee
Code 53
Department of Mathematics
Naval Postgraduate School
Monterey, CA 93940

Participation in the workshop is open. Registration Fee: $40 (includes copy of Proceedings).

Telephone inquiries may be made by calling (408) 646–2206 or 646–2622.

A better brand of calculus— e

Calculus with Analytic Geometry

ROBERT ELLIS and DENNY GULICK,
both of the University of Maryland at College Park

Calculus with Analytic Geometry is a complete basic textbook for the three-semester calculus sequence. The authors provide an honest but not overly rigorous treatment of all the essential topics of calculus: functions, limits, differentiation and integration in one and several variables, analytic geometry, series, and vector fields.

We think you will agree that we have a "better brand of calculus" when you see the straightforward motivation, the careful explanations, and the clear statement of definitions and theorems that together shape the unique foundation of this book. Other important elements of this new book include:

□ **a special emphasis on geometric intuition.**

□ **the extensive use of examples** to clarify new concepts and ideas.

□ **approximately 5,500 exercises** aimed at various levels of achievement.

□ **lively applications,** of which some of the more unusual examples are coughing, vascular branching, analysis of rainbows, mass of a binary star, dating of lunar rocks, and insulation of buildings.

□ **aids for the student** such as lists of key terms, expressions, and formulas, and review exercises, tables, and correct pronunciation for difficult mathematical terms.

□ **historical notes and comments on the "human side" of mathematics.**

□ **an attractive, open format** that features approximately 770 line drawings and important equations highlighted with color.

□ **an effective yet flexible organization** with supplementary material on differential equations and an Appendix with advanced proofs at the back of the book.

1060 pages (probable)
Publication: February 1978

A comprehensive *Solutions Manual*—with answers and *worked solutions* to all exercises in the textbook—will be available free of charge to instructors who adopt *Calculus with Analytic Geometry*. The manual will also be available for student purchase if the instructor desires.

Paperbound. 600 pages (probable)
Publication: February 1978

HBJ

Harcourt Brace Jovanovich, Inc.
New York • San Diego • Chicago • San Francisco • Atlanta

THE MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.
Washington, DC 20036

INDEX TO VOLUME 84, 1977
THE AMERICAN MATHEMATICAL MONTHLY

AUTHOR INDEX

MISCELLANEA

SUBJECT INDEX

This index uses the AMS (MOS) *Subject classification scheme* (1970), Math. Rev., 39 (1970) A1–A42 where a complete version can be found. Underlined entries indicate that the subject heading is secondary. In the (MOS) classification scheme the third level entries ending in 99 are "None of the above, but in this section." In this index the 99's carry the second level classification in which they lie.

PROBLEMS AND SOLUTIONS

PROBLEMS PROPOSED

Errata: E–1243, E–2440, E–2652 p. 567

PROBLEMS SOLVED

SOLUTIONS

Numbers in **boldface** type refer to problems, those in lightface to pages

See also p. 654

REVIEWS

Names of authors are in ordinary type, those of reviewers in capitals.

FILMS

TELEGRAPHIC REVIEWS

THE AMERICAN MATHEMATICAL MONTHLY

(FOUNDED IN 1894 BY BENJAMIN F. FINKEL)

AN OFFICIAL JOURNAL OF

THE MATHEMATICAL ASSOCIATION OF AMERICA

VOLUME 84 **NUMBER 10**

CODEN: AMMYAE

CONTENTS

DECEMBER **1977**

NOTICE TO AUTHORS

Research papers per se are unsuitable; see statement of policy (Vol. 81, p. 1).

Please follow the format in current issues of the MONTHLY. Manuscripts must be legibly typewritten or reproduced from typewritten copy, double spaced with wide margins and on one side of the paper. Three copies should be submitted to the appropriate editor and one kept by the author as protection against loss. The author's full address *must* appear at the end of the manuscript.

Backlog: Main Articles 16 months, Math. Notes 17 months, Research Problems 8 months, Classroom Notes 13 months, Math. Education 17 months.

EDITORIAL CORRESPONDENCE AND MAIN ARTICLES: to R. P. BOAS, Department of Mathematics, Northwestern University, Evanston, IL 60201; NOTES, etc.: to the corresponding Associate Editor; REPRINT PERMISSION: to LEONARD GILLMAN, Mathematical Association of America, University of Texas, Austin, Texas 78712 (see also the copyright notice below); ADVERTISING CORRESPONDENCE: to RAOUL HAILPERN, Mathematical Association of America, SUNY at Buffalo, Buffalo, N.Y. 14214; CHANGE OF ADDRESS and SUBSCRIPTIONS: to A. B. WILLCOX, Mathematical Association of America, 1225 Connecticut Ave., N. W., Washington, D.C. 20036; BACK ISSUES: Contact P. and H. Bliss Co., Middletown, CT 06457.

Annual dues for members of the Association, including a subscription to the American Mathematical Monthly, are $21.00 for each of the first two years of membership and $25.00 thereafter. Student Membership is available with annual dues of $15.00. For nonmembers the subscription price is $28.00.

PUBLISHED BY THE ASSOCIATION at Washington, D.C., and Menasha, Wisconsin, during the months of January, February, March, April, May, June-July, August-September, October, November, December.

Second-class postage paid at Washington, D.C., and additional mailing offices.

PRINTED IN THE UNITED STATES OF AMERICA

SINGULARITIES AND PLANE MAPS II: SKETCHING CATASTROPHES

JAMES CALLAHAN

1. Introduction.

Catastrophe is a phenomenon which appears in families of smooth real-valued functions. Roughly speaking, a catastrophe is an abrupt change in the number of minima possessed by functions in a family, as the parameters defining the family are varied continuously. The fundamental example is $f_c(x) = x^3 - cx$, which has one minimum when $c > 0$ and none when $c < 0$. It is important to consider minima apart from other kinds of critical points because the basis of many applications is a dynamical system endowed with a potential function V. The dynamic, in the form of the vector field $-\text{grad } V$, insures that the state of the system is found at one of the minima of V, because these are precisely the attractors of $-\text{grad } V$.(In fact, there is a more general form of catastrophe theory which is concerned with abrupt changes in the structure of the attractor sets of families of arbitrary, non-gradient dynamical systems.) If the state lies at a minimum of V which is destroyed through perturbation (as in Figure 1), the system will experience a catastrophic change as the state is directed by $-\text{grad } V$ to a

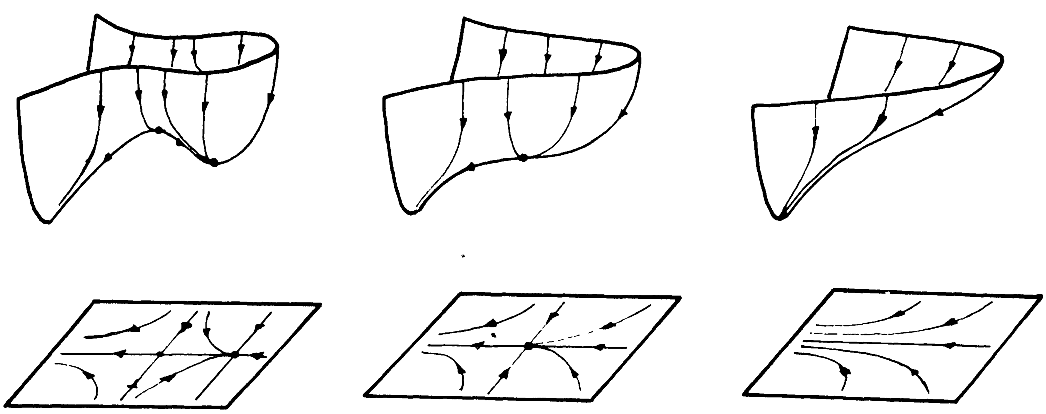


FIG. 1

new minimum. Because the parameters determine the nature of the flow and the number and location of the minima, they are often called *control variables.* Zeeman's catastrophe machine, with two controls, is a most vivid example [**13**, **23**]. In applications where it is the maxima which are important—for instance, in behavioral problems involving probability functions—it suffices to study the minima of the reciprocals of the functions being considered.

We shall say a function *undergoes catastrophe* when it is embedded in a family in such a way that the number of minima possessed by neighboring functions is not constant. Consider now a function f which has only non-degenerate critical points. It is a fundamental result of Morse theory (see [**6**], Theorems 3 and 4, and the references cited there) that all the critical points of any function near f will also be non-degenerate, and there will be precisely one, of the same index, near each critical point of f. Since the minima of a smooth function are precisely its critical points of index 0, it follows that any function near f has as many minima as f. In other words, a non-degenerate function cannot undergo catastrophe.

So we start instead with a function f having a degenerate critical point, and embed it in a family F_c. Assume that the controls $c = (c_1, \cdots, c_k)$ lie in a neighborhood of the origin in $\mathbb{R}^k$ and $F_0 = f$. Thom calls such an embedding an *unfolding*. The idea is that the degenerate critical point of f contains

within it the potential to "blossom," or unfold, into several non-degenerate critical points, which then appear in the nearby functions F_c. A **catastrophe** is thus a pair, consisting of

(a) the germ of a function f at a degenerate critical point;

(b) an unfolding of f, and the number of minima exhibited by the functions in the unfolding must vary.

It is important to realise that not every degenerate germ can be an element of a catastrophe. For example, $f(x, y) = x^4 - y^2$ has a degenerate saddle at the origin, but no function in any unfolding of f can have a minimum. Thus, f does not undergo catastrophe.

Let us look more closely at the functions in a particular unfolding. In general, there will always be some functions which have a degenerate germ which can itself undergo catastrophe, but these germs will always be "less degenerate" than the germ at the center of the catastrophe. This is illustrated by the familiar cusp catastrophe, when its germ $y = x^4$ is unfolded by the family $F_{a,b}(x) = x^4 - ax - bx^2$ (cf. Figure 13 in [**6**]). One can readily check that the critical points of the function $F_{a,b}$ are all non-degenerate, unless the control point (a, b) lies on the cusp-shaped curve $27a^2 = 8b^3$. For each point $(a, b) \neq (0,0)$ on this curve, $F_{a,b}$ has a degenerate germ equivalent to $y = x^3$. This germ undergoes catastrophe, for example when it is embedded in the family $G_c(x) = x^3 + cx : G_c$ has one minimum if $c < 0$ but none if $c > 0$. We consider $y = x^3$ "less degenerate" than $y = x^4$ in the sense that all the changes which $y = x^3$ can undergo are adequately exhibited by a one-parameter family, while two parameters are needed for $y = x^4$. Thus, in the control space of the unfolding $F_{a,b}$, there are three kinds of points: (a) the origin, corresponding to the catastrophe $y = x^4$; (b) the other points of the cusp curve $27a^2 = 8b^3$, corresponding to the less degenerate catastrophe $y = x^3$ which "evolves" from $y = x^4$; (c) points not on the cusp curve, corresponding to non-catastrophic functions. The cusp curve is called the catastrophe set of $y = x^4$; its structure reveals much about the catastrophe. Following this pattern, we define the **catastrophe set** K of an arbitrary catastrophe (f, F) to consist of those points c in the control space $\mathbb{R}^k$ for which the function F_c has a degenerate germ which itself can undergo catastrophe. The points of K will correspond to catastrophes less degenerate than (f, F), and the structure of K will give insights about the relations between catastrophes.

Here now are two of the basic questions in catastrophe theory which we shall consider:

I CLASSIFICATION. What catastrophes can occur? The aim is a list of degenerate germs and unfoldings which is complete in some useful sense.

II GEOMETRY. How are the different catastrophes related? The points in a given catastrophe set K correspond to other, less potent, catastrophes. What is the geometry of K, and what other catastrophes appear in it?

There is yet no full answer to these questions. We shall describe only a special class of catastrophes, involving what Arnold calls the *simple* germs. In this case there is a complete classification and a reasonable description of the geometry. A good deal is known about the more general case, but we shall not go into it. However, see [**3**, **4**, **14**, **15**, **24**].

We shall assume as background the definitions and results of [**6**], together with the excellent introduction to catastrophe theory provided by E. C. Zeeman in *Scientific American* [**22**]. Of course the origin of catastrophe theory is in the work of René Thom. The geometric questions we shall consider were first treated in Thom's book [**16**, chapter 5], and also in [**5**, **8**, **20**]. Extensive bibliographies of the whole subject can be found in [**4**, **17**].

2. Classification of Catastrophes.

The first step in producing a manageable list of catastrophes is to lump together equivalent ones. Altogether, three different notions of equivalence will arise, each associated with a particular class of objects.

(a) Two *maps*, or *map germs*, $f, g : N \to P$ of *arbitrary smooth manifolds* are equivalent if there are diffeomorphisms, or diffeomorphism germs, $h : N \to N$ and $k : P \to P$ for which $g = k \circ f \circ h$. This is called **left-right equivalence**; it is the basic notion, from which the following two are derived.

(b) Two *real-valued functions* $f, g : \mathbb{R}^n \to \mathbb{R}$ are equivalent if they are left-right equivalent, as in (a), and k is orientation preserving. The additional condition is needed to prevent f and $-f$ from being equivalent; these functions have, in general, a different number of minima and must be considered distinct in catastrophe theory.

(c) Two *real-valued germs* $f, g : \mathbb{R}^n \to \mathbb{R}$ are equivalent if there is a diffeomorphism germ $h : \mathbb{R}^n \to \mathbb{R}^n$ for which $g = f \circ h$. This is called **right-equivalence**. Thus, two catastrophes (f, F) and (g, G) will be equivalent if the germs f and g are right-equivalent and F and G are equivalent as functions in the sense (b).

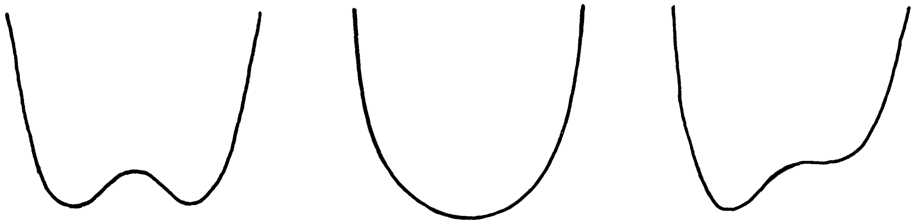
The next step is to eliminate the non-simple germs. A germ is **simple** if the functions in any unfolding of it represent only a finite number of inequivalent types. This obviously simplifies the study of the catastrophe set K. Non-simple germs are quite common. In fact, no germ which is degenerate in three or more variables is simple. More precisely, suppose r denotes the rank of the Hessian of a germ f, and $r^* = n - r$ is its **corank**. Then by definition f is degenerate if $r^* \geqq 1$; by a theorem **[2]**, it is non-simple if $r^* \geqq 3$. The argument may be summarized in the following way. Suppose, for the sake of illustration, that f and g are three-variable germs, each with a corank 3 singularity at the origin. They are right-equivalent only if there is a diffeomorphism germ $h : \mathbb{R}^3, 0 \to \mathbb{R}^3, 0$ for which $g = f \circ h$. Let $f = f_3 +$ higher order terms, where f_3 is a homogeneous cubic polynomial, and similarly for g, and let $h = h_1 +$ higher terms, where $h_1 : \mathbb{R}^3 \to \mathbb{R}^3$ is an invertible linear map. Then $g = f \circ h$ implies $g_3 = f_3 \circ h_1$; this reduces the question to the equivalence of cubic forms under linear coordinate changes. However, the space of cubic forms has dimension 10, while the linear group $GL(3, \mathbb{R})$ has only dimension 9. Thus there is at least a one-parameter family of inequivalent forms (such a parameter is called a *modulus*), and hence a one-parameter family of inequivalent corank 3 germs. (Not every two-variable germ is simple, however. The family $G_c = x^4 + y^4 + cx^2y^2$ unfolding $g = x^4 + y^4$ also consists of pairwise inequivalent functions. The argument is similar; the dimension of the space of quartic forms in two variables is 5, but $\dim GL(2, \mathbb{R}) = 4$. See also **[3, 14]**.)

When the non-simple catastrophes are eliminated, we can deal exclusively with two-variable germs and suffer no loss of generality. For suppose we consider any simple germ $f^*(x_1, \cdots, x_n)$ undergoing catastrophe. Since corank $f^* \leqq 2$, a suitable coordinate change makes f^* equivalent to a germ of the form $f(x_1, x_2) + x_3^2 + \cdots + x_n^2$. (This result can be viewed as a generalization of the Morse lemma **[2, 7, 18]**.)

Moreover, any unfolding of f^* is equivalent to $F^*(x_1, \ldots, x_n) = F(x_1, x_2) + x_3^2 + \cdots + x_n^2$, where F is an unfolding of f. The only critical points of F^* have the form $(a_1, a_2, 0, \ldots, 0)$, where (a_1, a_2) is a critical point of the same kind (i.e., degenerate, or non-degenerate, of index j) for F. Consequently, all changes experienced by functions near f^* can be found in the unfolding $F(x_1, x_2)$. In the other direction, suppose $\varphi(x)$ is a one-variable simple germ. Enlarging it to $f(x_1, x_2) = \varphi(x_1) + x_2^2$ and unfolding f likewise exhibits everything that can happen near φ.

The last step is to consider the possible unfoldings of a given degenerate simple germ. The difficulty here is that a particular unfolding may be incomplete in that it fails to include all possible nearby types. Take, for example, the unfolding $H_b(x) = x^4 - bx^2$ of $y = x^4$. It includes no functions equivalent to $G_a(x) = x^4 - ax - \frac{3}{2}a^{\frac{2}{3}}x^2$, although these are near the function $y = x^4$. (The coefficient $\frac{3}{2}a^{\frac{2}{3}}$ was chosen to ensure that G_a had an inflection point.) See Figure 2. Likewise, G_a is an incomplete unfolding of $y = x^4$. These faults are remedied by $F_{a,b}(x) = x^4 - ax - bx^2$, which incorporates both H and G and which, furthermore, includes a function equivalent to every one near $y = x^4$. An unfolding which has this last property is said to be **universal**. Obviously, in making a list of catastrophe pairs (f, F), it suffices to let F be a universal unfolding of f.

Up until now there has been a confusing interplay between germs and functions. We began with a



H_b; $b>0$ $H_0=G_0$ G_a; $a>0$

FIG. 2

function, but seeing that the cause of catastrophe was local, we switched attention to its *germ* at a degenerate critical point. With classification in mind, we lumped together different functions undergoing catastrophe if their degenerate germs were right-equivalent. Turning to unfoldings, though, we went back to *functions* in order to see how a degenerate critical point could blossom into less degenerate, and finally non-degenerate, critical points. A germ, because it is attached to a single point, cannot provide this kind of information. For example, the only *germs* near $y=x^4$ are $y=x^3$, $y=\pm x^2$, and $y=x$ (and others right-equivalent to these). But *functions* near $y=x^4$ include those equivalent to

$y=x^2$ one minimum

$y=x^4+x^3$ one minimum and one horizontal inflection

$y=x^4-x^2$ two minima and one maximum.

It is the latter information which we are after and which is important in applications. (Note: two functions $f, g:\mathbb{R}^n\to\mathbb{R}$ are neighbors in the sense of the Whitney C^∞-topology on $C^\infty(\mathbb{R}^n,\mathbb{R})$ (see [**5**] or [**8**]).)

We can eliminate some of the awkwardness of this dual approach, and recast the discussion entirely in the language of germs. The trick is to modify slightly the definition of an unfolding so that the control variables appear on an equal footing with the rest. Henceforth an **unfolding** of a germ $f:\mathbb{R}^n,0\to\mathbb{R},0$ is a smooth germ

$$F:\mathbb{R}^k\times\mathbb{R}^n,0\to\mathbb{R},0:\ (c,x)\to F(c,x)$$

for which $F(0,x)=f(x)$. Here c is a **control point**, $\mathbb{R}^k$ is the **control space**, and k is the dimension of the unfolding.

An immediate benefit of this modification is that we can use Mather's procedure for constructing a universal unfolding [**11**, **12**, **18**]. Let $\mathscr{E}$ denote the local ring of germs $g:\mathbb{R}^n,0\to\mathbb{R}$ and let m be its maximal ideal, consisting of all germs vanishing at the origin. It can be shown that m is finitely generated, by the n coordinate functions $x_i:\mathbb{R}^n,0\to\mathbb{R},0$. Both $\mathscr{E}$ and m are infinite-dimensional real vector spaces.

Suppose f is a germ to be unfolded. Form the **Jacobian ideal** J_f of f; this is generated by the germs of $\partial f/\partial x_1,\ldots,\partial f/\partial x_n$ in $\mathscr{E}$. These germs all vanish at the origin, because we are assuming f has a (degenerate) critical point there. Thus $J_f\subset m$ and we can consider the quotient m/J_f. It is a real vector space, and is finite-dimensional for all the simple germs f [**18**]. Its dimension k is called the **codimension** of f (because m/J_f is a complement to the tangent space of the right-equivalence class of f in m). Finally, if the germs $u_1,\ldots,u_k:\mathbb{R}^n,0\to\mathbb{R},0$ form a basis for m/J_f as a real vector space, then

$$F(c_1,\ldots,c_k,x)=f(x)+c_1u_1(x)+\cdots+c_ku_k(x)$$

gives a universal unfolding of the simple germ f. Obviously a universal unfolding is not unique. However, any one defined this way at least has minimal dimension. That is, an unfolding cannot be universal if its dimension is less than the codimension of the germ being unfolded.

As an example, consider $f = x^4 + y^2$. Then $n = 2$ so $m = (x, y)$ and $J_f = (x^3, y)$. A basis for $m/J_f = (x, y)/(x^3, y)$ is $\{x, x^2\}$ and it leads to the unfolding

$$F(a, b, x, y) = x^4 + y^2 + ax + bx^2.$$

A slightly more complicated example is $f = x^3 + xy^3$, which is E_7 in the list below. This time $J_f = (3x^2 + y^3, xy^2)$ and one possible basis for $m/J_f = (x, y)/(3x^2 + y^3, xy^2)$ is $\{x, y, x^2, xy, y^2, x^2y\}$.

Table 1 is a complete list of the simple germs which undergo catastrophe; k can take any positive integer value. The list was worked out by Arnold in [2], by a straightforward consideration of cases. The A, D, E notation for simple germs was introduced by Arnold in the same paper; he was struck by similarities between the blossoms of these germs (for a definition of the term, see Section 3 and Table 3, below) and the Coxeter–Dynkin diagrams of the simple Lie groups A_n, D_n, E_n. There are further coincidences with other mathematical classification problems, for example, the finite subgroups of the orthogonal group $O(3, \mathbb{R})$ and hence the regular n-gons and Platonic solids. Arnold's paper [2] is thoroughly readable and is an excellent introduction to all these ideas. His later papers [3, 4] (and see also the bibliography in [4]) show how this viewpoint has been developed to analyze even the non-simple germs.

TABLE 1.

Name	Symbol	Germ
Cuspoid	A_{k+1}	$x^{k+2} + y^2$
Dual cuspoid	A^*_{2k+1}	$-x^{2k+2} + y^2$
Hyperbolic umbilic	D_{2k+2}	$x(x^{2k} + y^2)$
Elliptic umbilic	D^*_{2k+2}	$x(-x^{2k} + y^2)$
Parabolic umbilics	D_{2k+3}	$x(x^{2k+1} + y^2)$
	$-D_{2k+3}$	$-x(x^{2k+1} + y^2)$
Exceptional	E_6	$x^3 + y^4$
	$-E_6$	$x^3 - y^4$
	E_7	$x^3 + xy^3$
	E_8	$x^3 + y^5$

For purposes of discussion, we supplement this list with Table 2. These germs are not catastrophes; either they are non-degenerate (and hence stable) or else they have no minima in their unfoldings.

TABLE 2

Name	Symbol	Germ
Stable minimum	A_1	$x^2 + y^2$
Stable saddle	A^*_1	$-x^2 + y^2$
Stable maximum	$-A_1$	$-x^2 - y^2$
—	$-A_{k+1}$	$-x^{k+2} - v^2$
—	$-A^*_{2k+1}$	$x^{2k+2} - y^2$

A final remark about the subscript appearing in a symbol: if P stands for any one of $\pm A$, $\pm A^*$, $\pm D$, D^*, or $\pm E$, then in every case the germ P_j has codimension $j - 1$ and Milnor number j. The **Milnor number** of a germ is the maximum number of critical points which can appear in an unfolding.

3. Geometry: Introduction and examples.

We shall consider two complementary approaches to aspects of catastrophe theory. Each is suited

to answer one or the other of the questions raised in the introduction. The first is due to Arnold [**2**, **3**], and uses Picard–Lefschetz theory and Coxeter–Dynkin diagrams to decide which other catastrophes occur in an unfolding of a given germ. Actually, Arnold's methods apply to the complex case; the adaptation to real functions which we shall follow has been worked out by A'Campo [**1**] and Gusein-Zade [**9**]. In the second approach, Zeeman [**18**, **21**, **22**, **24**] represents each catastrophe set K as the apparent contour of a map, the catastrophe map χ, associated with an unfolding. (The **apparent contour** of a smooth map $T: N \to P$ is the image of the singular set of T in P.) Knowing that K is an apparent contour lends insight to its geometry. A couple of examples will illustrate both methods and show how they are linked.

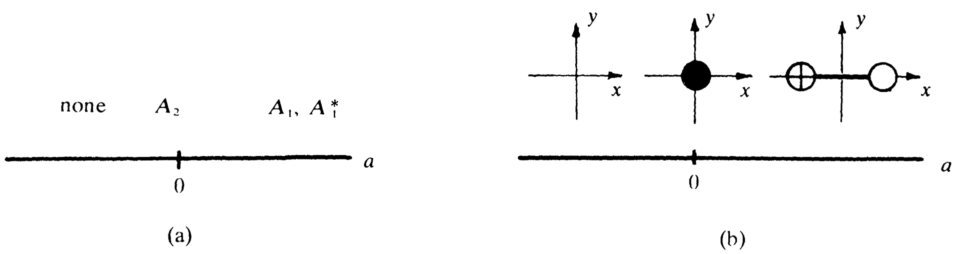


FIG. 3

The basic example is $A_2: f(x, y) = x^3 + y^2$. Using Mather's procedure we select the particular universal unfolding $F(a, x, y) = x^3 + y^2 - ax$. Figure 3a lists the kinds of critical points which appear in $F(a, -, -)$ as a varies in the one-dimensional control space. Note that a catastrophe does indeed occur, because there is an A_1-type point—i.e., a minimum—present when a is positive but none when a is negative. In Figure 3b we find roughly the same information about F but it is presented in a more graphic fashion. Namely, the actual configuration of the critical points of $F(a, -, -)$ in the xy-plane is displayed, for each of the cases $a < 0$, $a = 0$, and $a > 0$. We are using the following scheme, here and in the sequel, to distinguish between the critical points of an unfolding function F defined on the xy-plane:

$$\text{non-degenerate} \begin{cases} \oplus & \text{minimum} \\ \bigcirc & \text{saddle} \\ \ominus & \text{maximum} \end{cases}$$

$$\text{degenerate} \quad \bullet \quad \text{(any type).}$$

We shall also connect two critical points by a line, as in the case $a > 0$ above, whenever there is a trajectory of the gradient dynamic $-\operatorname{grad} F$ linking them. (The notation for the non-degenerate critical points is suggested by the signature of the quadratic form corresponding to each (see Table 2). The signature of any quadratic form is defined as the number of positive eigenvalues minus the number of negative eigenvalues. Thus, the signatures of $x^2 + y^2$ (minimum), $x^2 - y^2$ (saddle), $-x^2 - y^2$ (maximum) are $+2$, 0, -2, respectively.)

Figure 3b gives us the following interpretation of the A_2 catastrophe: a saddle $\bigcirc$ and a minimum $\oplus$ coalesce momentarily as $\bullet$ and then annihilate one another as the control a decreases and passes through the origin. They are created again as a is varied in the other direction. Because of this interpretation, A_2 is sometimes called the *birth-death* catastrophe. It is also called the *saddle-node* catastrophe, because the minimum $\oplus$ appears as a node in the gradient vector field. Notice that only the *relative* positions of the critical points are needed to detect catastrophe. For this reason the xy-plane is eliminated from all subsequent illustrations; only the configuration of the critical points is shown.

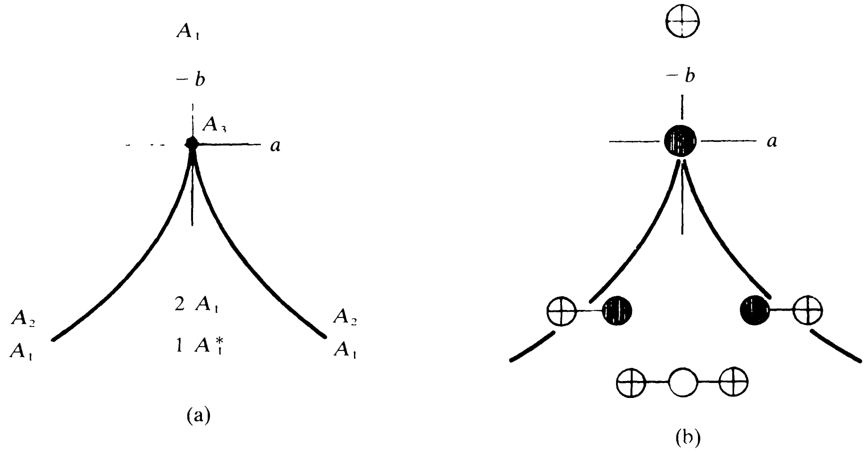


FIG. 4

Figure 4 is the analogue of Figure 3 for the cusp A_3, unfolded as $f(a, b, x, y) = x^4 + y^2 - ax - bx^2$. The catastrophe set K is the curve $27a^2 = 8b^3$, because at each of these points the corresponding function has a degenerate germ, namely either A_2 or A_3. However, the usefulness of the symbolic diagram (b) begins to be more apparent. For example, the arrangement of the critical points inside the cusp makes it clear that two different saddle-node catastrophes can occur, as the saddle collides with the minimum on either its right or its left. Both collisions do occur, and produce the two sides of the cusp.

We can read this diagram another way. First, let us call any configuration—like those appearing in Figures 3b and 4b—which consists of the maximum possible number of critical points of an unfolding function of a catastrophe a **blossom** of that catastrophe. An individual symbol $\oplus$, $\bigcirc$, $\ominus$ in a blossom will be called a petal. (The number of petals in a blossom must therefore be equal to the Milnor number of the germ.) Thus, a blossom for A_3 is $\oplus$—$\bigcirc$—$\oplus$ and, for A_2, $\bigcirc$—$\oplus$ (or $\oplus$—$\bigcirc$, which appears if A_2 is represented by the equivalent germ $-x^3 + y^2$). Now suppose we follow a path in Figure 4b clockwise from the top. We meet, in succession, the symbols $\oplus$, $\bullet$—$\oplus$, $\oplus$—$\bigcirc$—$\oplus$. The petal $\oplus$ on the right represents a minimum which is not disturbed during the control changes along the path; delete it. What results is the sequence (blank), $\bullet$, $\oplus$—$\bigcirc$ of A_2. In other words, we have found once again that A_2 lies in (the catastrophe set of) A_3. But there is no need actually to traverse the path in the control space; we can infer that the A_2 sequence must appear simply by removing a $\oplus$ petal from the A_3 blossom.

In general, if a blossom of P_{k-1} appears when a petal is removed from a blossom of P_k, then P_{k-1} appears in the catastrophe set of P_k. By downward induction on the Milnor number k (i.e., by plucking further petals), we can discover all catastrophes appearing in a given catastrophe set. So we need only determine which P_{k-1} are found in P_k.

This is not magic. Removing a petal from a blossom is to be interpreted simply as restricting the control variables so that the presence and non-degeneracy of the critical point represented by the petal are never destroyed. The restriction essentially decreases the number of control variables by one (e.g., we chose a one-dimensional path in the two-dimensional control space of A_3 to find A_2), and this reduction in the dimension of the control space means we are dealing with a (universal) unfolding of a new catastrophe whose codimension is one less than the one we started with.

This method of Arnold will tell us what catastrophes appear in a particular unfolding, but not where. That information is supplied by Zeeman's catastrophe manifold M. Essentially, M is the graph of the blossom of an unfolding as a function of the controls. In more precise terms, the **catastrophe manifold** M of an unfolding F is the locus

$$\left\{(c,x):\frac{\partial F}{\partial x_1}(c,x)=\cdots=\frac{\partial F}{\partial x_n}(c,x)=0\right\}\subset\mathbb{R}^k\times\mathbb{R}^n.$$

Since it is given by n conditions in a $(k+n)$-dimensional space, M has dimension k, which is equal to the dimension of the control space (and the unfolding). The informal definition amounts to the same thing, because a blossom is just the set of critical points of an unfolding function.

The easiest way to see how M gives information about the structure of a catastrophe set is to look at some examples. Figure 5 gives M for A_2 and the unfolding we used above. It is a "two-sheeted"

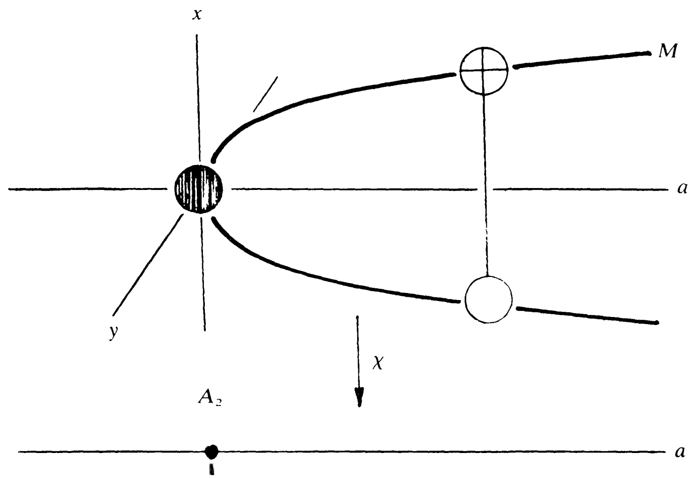


FIG. 5

curve over the a-axis, folded at the origin (hence the usual name for A_2: the **fold**). Figure 6 does the same thing for A_3. Notice, however, that the variable y has been suppressed to make it possible to draw a picture. This causes no problems because the y-coordinate of every critical point is 0. Thus the picture, which represents the subspace $y=0$ of the full unfolding space $\mathbb{R}^2\times\mathbb{R}^2$, misses nothing essential. We could have eliminated the y-direction from Figure 5 as well. More generally, the y-coordinate of any corank 1 germ (A or A^*) can be ignored.

Define the **catastrophe map** $\chi:M\to\mathbb{R}^k$ to be the restriction of the projection $\pi:\mathbb{R}^k\times\mathbb{R}^n\to\mathbb{R}^k$ to M. It is evident from the figures, and in any case it is not difficult to prove, that the singular set of χ (in M) consists of the degenerate critical points of the unfolding (these are the solid circles in the diagram) and the projection of the singular set to the control space is the catastrophe set K. In other words, K is the apparent contour of χ.

Notice that the dimension of the set of A_2 points is 1 in Figure 6 but it is 0 in Figure 5, while in both cases the set has codimension 1 in the control space. It is generally true that the dimension, but not the codimension, of the subset of points of a given type in a catastrophe set can vary from one example to another, This leads to a complication in relating the catastrophe sets of two different unfoldings of a given germ. For example, $H(a,b,x)=x^3-ax-b$ is another universal unfolding of $y=x^3$, but M and K for H (Fig. 7) are each one dimension larger than the corresponding sets for the unfolding F of Figure 5. In general, if F_1 and F_2 are universal unfoldings of f, of dimensions k_1 and k_2 respectively, and $j=k_1-k_2\geqq 0$, then we obtain the following commutative diagram. The vertical lines represent diffeomorphisms.

$$\begin{array}{ccccc}
M_1 & \xrightarrow{\chi_1} & \mathbb{R}^{k_1} & \supset & K_1\\
| & & | & & |\\
M_2\times\mathbb{R}^j & \xrightarrow[\chi_2\times id]{} & \mathbb{R}^{k_2}\times\mathbb{R}^j & \supset & K_2\times\mathbb{R}^j
\end{array}$$

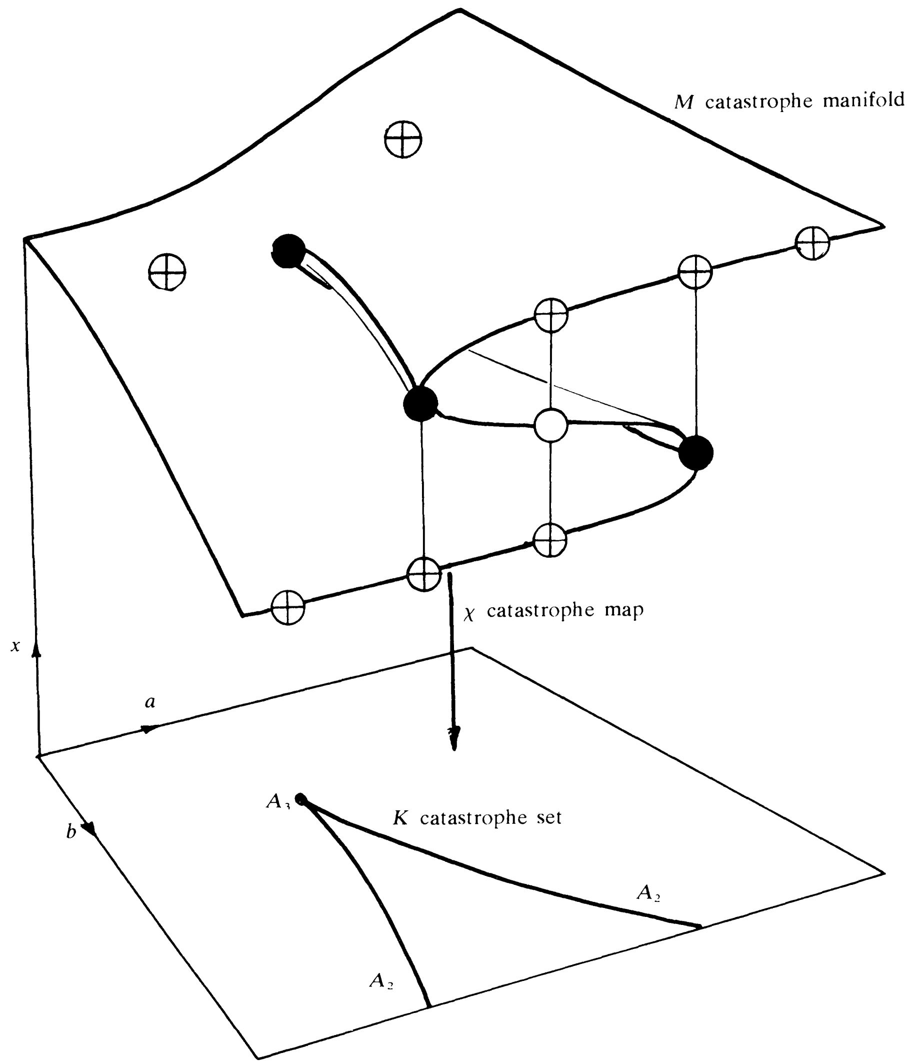


FIG. 6

Two final provisos: First, since an unfolding F is only a germ, the same is true of its catastrophe manifold M and map χ. In Zeeman's approach, this **catastrophe germ** $\chi : M, 0 \to \mathbb{R}^k, 0$ is the basic object of study. Second, it will appear when we come to the umbilics that certain components of the singular set of χ consist of degenerate critical points which do *not* undergo catastrophe (because their blossoms contain only saddles and maxima). We must modify the claim that K is the apparent contour of χ by excluding from K the images of these non-catastrophic components.

4. Arnold geometry.

In this section we obtain the blossoms of the simple catastrophes and use this information to determine the constituents of the different catastrophe sets. In each case we must construct an unfolding, choose a function having the maximum number of critical points, and then determine the relative positions of those critical points. It turns out to be particularly advantageous to choose a function whose saddles are all at the same level. The reason is this: by sketching the level curve of such a function at the level of the saddles, we determine at least the positions of the saddles—they appear

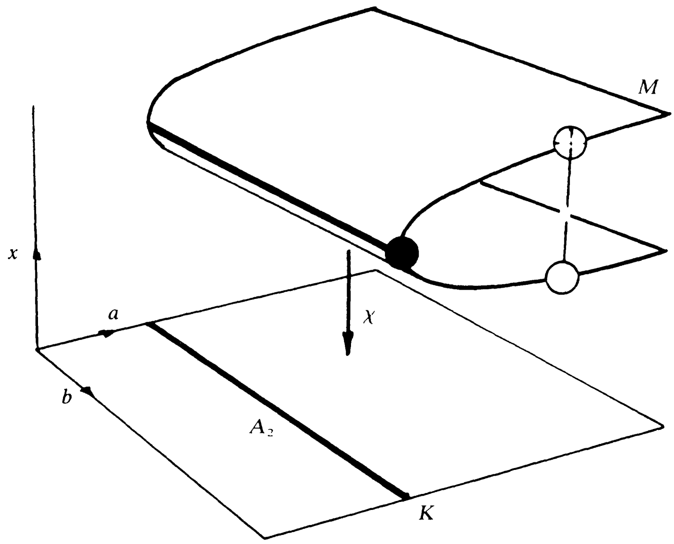


FIG. 7

as the double points of the level curve. Because the functions we deal with have a rather simple form, we can then infer the positions of the other critical points, that is, the maxima and minima. Working through the details should make this clearer.

The cuspoids and their duals. All the germs A, A^* (and $-A$, which we include for the sake of completeness) can be treated similarly. Let $(A)_{m+1} = \pm x^{m+2} \pm y^2$ represent any one of them; we unfold it by

$$F(c_1, \ldots, c_m, x, y) = (A)_{m+1} - c_1 x - \cdots - c_m x^m, m \geqq 1.$$

However, to select an appropriate function and obtain the blossom it will be more convenient to consider individual cases.

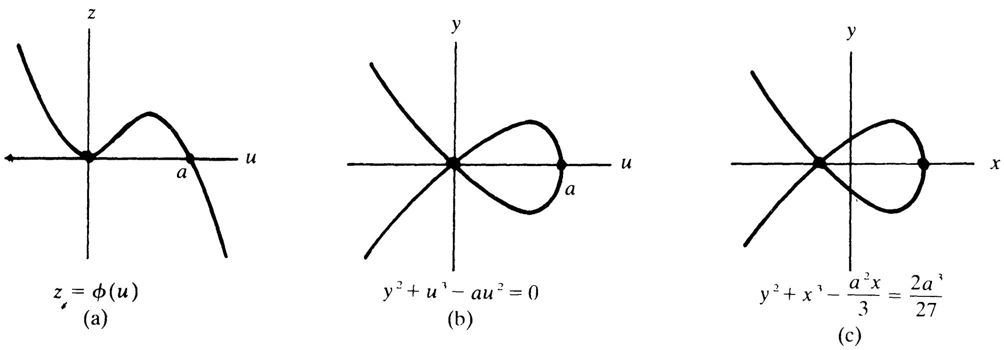


FIG. 8

A_2: We have to analyze the unfolding function $F(c, x, y) = y^2 + x^3 - cx$. This is not difficult to do directly, but we choose instead a more elaborate routine which carries over to the other cuspoids. Consider the auxiliary functions $\phi(u) = u^2(a-u)$ (Fig. 8a) and $F^*(u, y) = y^2 - \phi(u)$, $a > 0$. The given unfolding F is related to F^* through the translation $u = x + a/3$, as follows:

$$F^*(u, y) = y^2 - \phi\left(x + \frac{a}{3}\right) = y^2 + x^3 - \frac{a^2}{3}x - \frac{2a^3}{27} = F\left(\frac{a^2}{3}, x, y\right) - \frac{2a^3}{27}.$$

Now F^* has a saddle at the origin, and the saddle appears as a double point of the zero level curve of F^* in Figure 8b. Hence the unfolding function $F(a^2/3, -, -)$ has a saddle at $(-a/3, 0)$; its $2a^3/27$-level is a translate of the zero-level of F^* and is shown in Figure 8c. The closed loop of each of these level curves must contain a critical point, and in fact F has a minimum at $(a/3, 0)$. Because there is a trajectory of $-\operatorname{grad} F$ which flows from the saddle to the minimum, the ○ and ⊕ must be connected by a line, and the blossom of A_2 is ○—⊕.

A_{2k}: Following the pattern of A_2, start with

$$\phi(u) = (ka - u) \prod_{j=1}^{k-1} (ja - u)^2$$

and the graph $y^2 - \phi(u) = 0$ (Figures 9a and b). As before, a translation eliminates the unwanted term

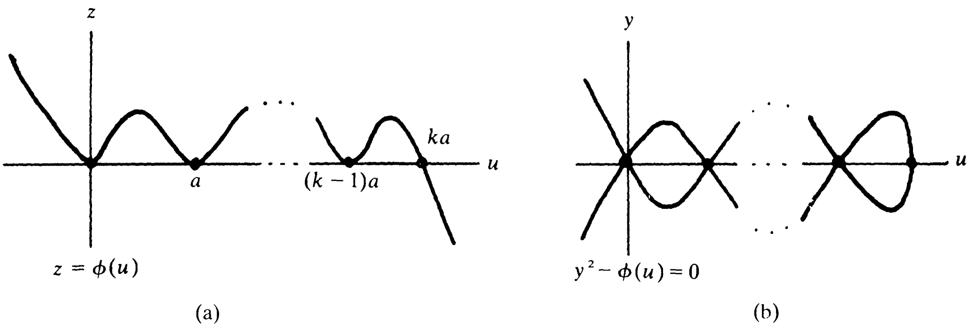


FIG. 9

of degree $2k$. What results is the level curve

$$y^2 - \phi\left(x + \frac{k^2 a}{2k+1}\right) = \phi\left(\frac{k^2 a}{2k+1}\right)$$

of a function in our unfolding of A_{2k}. Of course, since a translation will always put $y^2 - \phi$ into the proper form, we need not actually carry it out. The blossom of A_{2k} can be read directly from Figure 9b; it is ○—⊕—···—○—⊕ (k minima).

A_{2k+1}: If $k = 2q$, use the function

$$\phi_{2q}(x) = ((q+1)^2 a^2 - x^2) \prod_{j=1}^{q} (j^2 a^2 - x^2)^2,$$

and if $k = 2q + 1$, use

$$\phi_{2q+1}(x) = x^2 \phi_{2q}(x).$$

(Let $\phi_0(x) = a^2 - x^2$; then the stable germ A_1 will be included in this analysis.) Since these are all even functions, they contain no unwanted term of degree $2k + 1$, so translation is not necessary. Figure 10 illustrates the cases $k = 1, 2$; others are similar. The blossom of A_{2k+1} is ⊕—○—···—○—⊕ ($k + 1$ minima).

A^*_{2k+1}: If $k = 2q - 1$, use the function

$$\phi_{2q-1}(x) = \prod_{j=1}^{q} (x^2 - j^2 a^2)^2$$

and if $k = 2q$, use

$$\phi_{2q}(x) = x^2 \phi_{2q+1}(x).$$

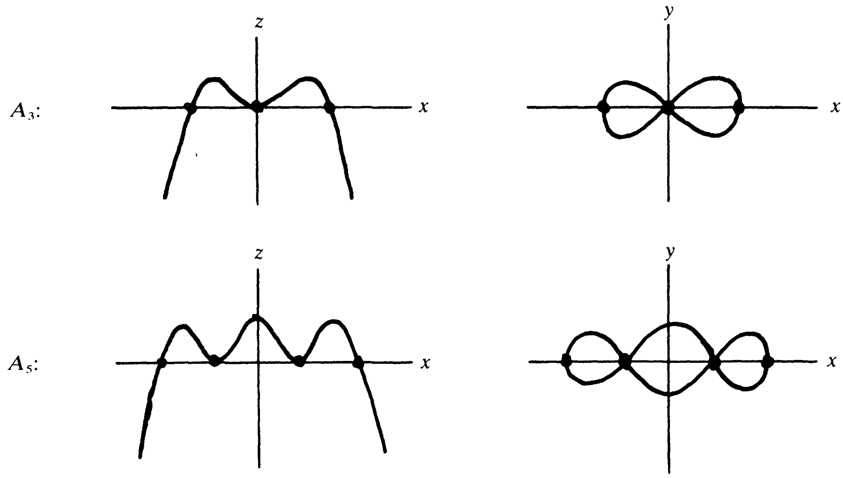


FIG. 10

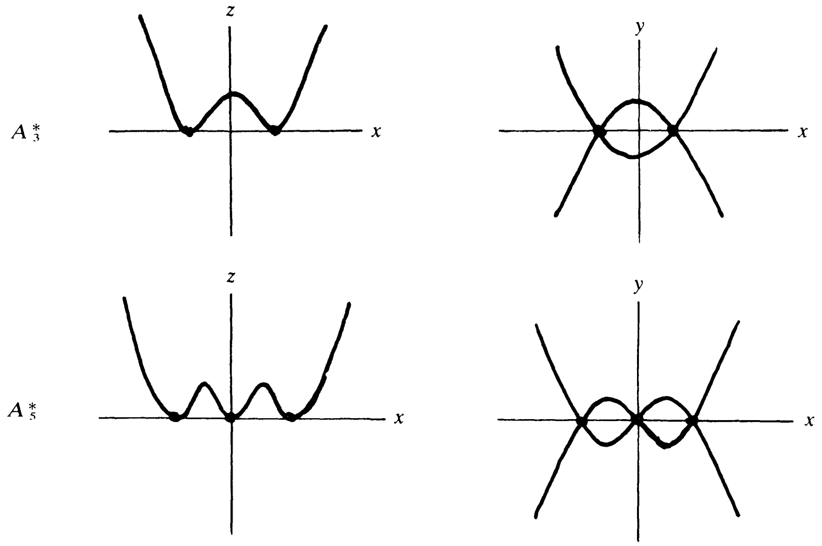


FIG. 11

The blossom of A^*_{2k+1} is ○—⊕—···—⊕—○ (k minima). See Figure 11.

$-A_m$ and $-A^*_m$: If F is an unfolding for A_m or A^*_m, then $-F$ is for $-A_m$ or $-A^*_m$; minima and maxima are interchanged. Thus:

$-A_{2k}$: ○—⊖—···—○—⊖

$-A_{2k+1}$: ⊖—○—···—○—⊖

$-A^*_{2k+1}$: ○—⊖—···—⊖—○ .

Obviously none of these is a catastrophe germ.

The umbilics. Let $(D)_{m+2}$, $m \geqq 2$ represent an arbitrary umbilic. We shall use the unfolding

$$H(a, b, c_1, c_2, \cdots, c_{m-1}, x, y) = (D)_{m+2} - (ax + by + c_1 y^2 + c_2 x^2 + \cdots + c_{m-1} x^{m-1}).$$

Observe that $(D)_{m+2} = x(A)_{m-1}$; we shall exploit this relationship to determine the blossoms of the umbilics.

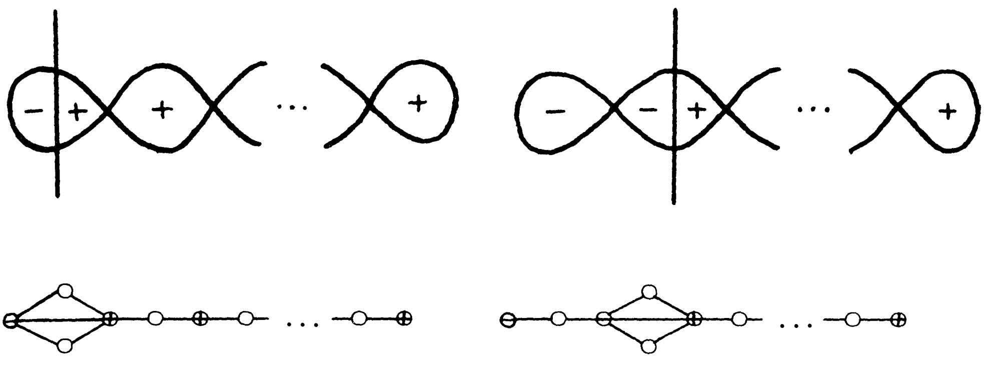
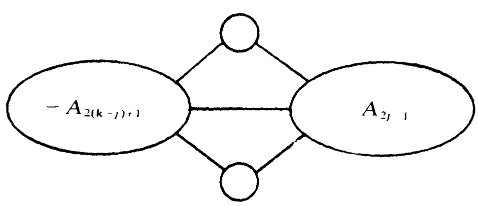
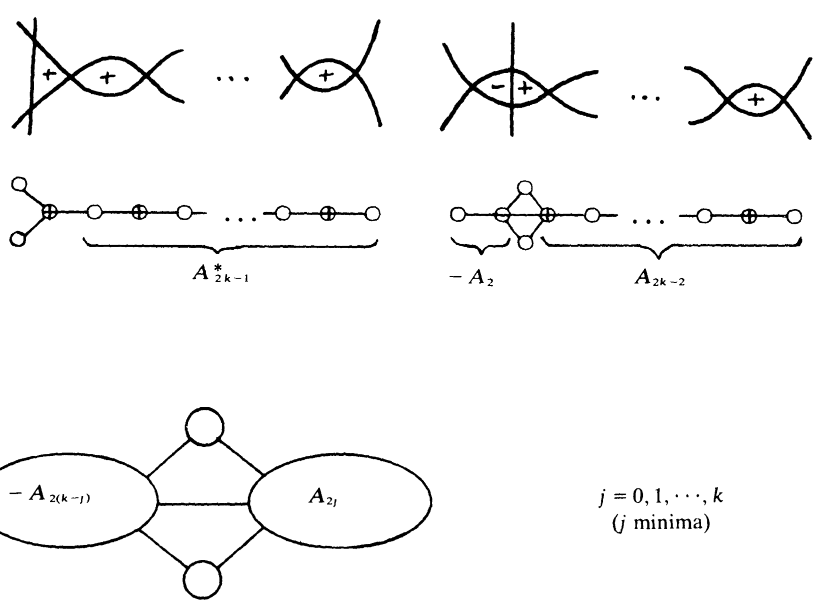


FIG. 12



$j = 1, \ldots, k$
(j minima)

FIG. 13—D_{2k+2}



$j = 0, 1, \cdots, k$
(j minima)

FIG. 14—D^*_{2k+2}

$\boldsymbol{D_{2k+2}}$: Let $y^2 - \phi(x)$ be the function selected above from the unfolding of A_{2k-1}. Then $\Phi(x, y) = (x - e)[y^2 - \phi(x)]$ (or a translate of it parallel to the x-axis) will be found in the unfolding

just given for D_{2k+2}. The zero-level of Φ is the union of the k-looped level curve of A_{2k-1} and a vertical line $x = e$, which can be made to pass through any loop by varying e. Figure 12 illustrates two possibilities, each leading to a blossom. In fact, there are k different blossoms, illustrated schematically in Figure 13.

$\boldsymbol{D^*_{2k+2}}$: Start with the function $y^2 - \phi(x)$ used to determine the blossom of A^*_{2k-1} and proceed as with the hyperbolic umbilics. Figure 14 illustrates two possibilities and a schematic diagram of the $k+1$ different blossoms. Interpret $\pm A_0$ as the empty set.

$\pm \boldsymbol{D_{2k+3}}$: In both cases the unfolding function for A_{2k} is involved. Two blossoms for D_{2k+3}, together with the schematic diagrams for D_{2k+3} and $-D_{2k+3}$, are given in Figure 15. Notice that the

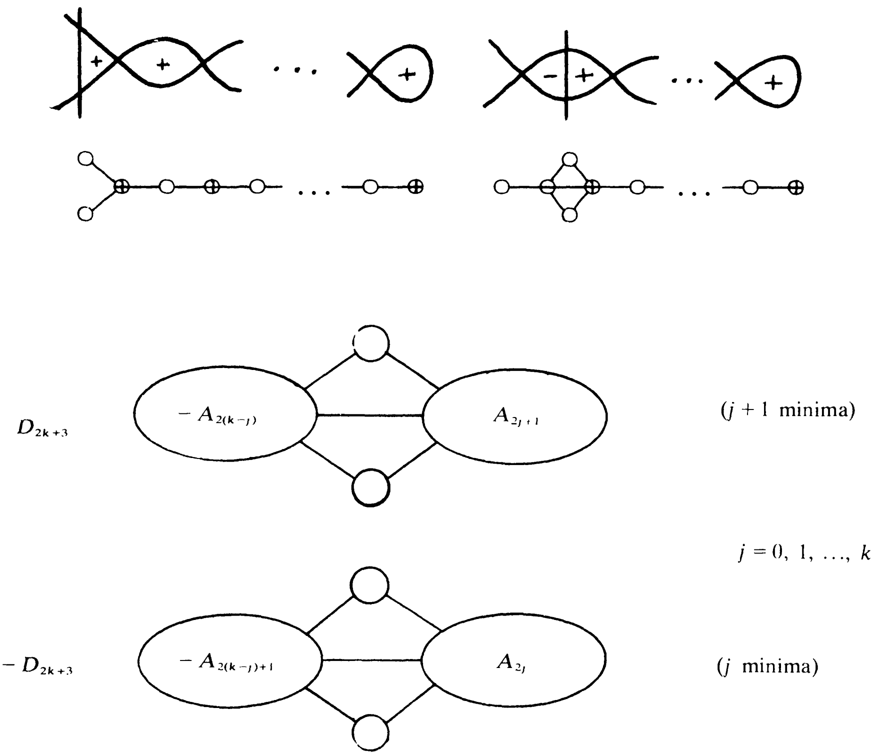


FIG. 15

parabolic forms are hybrids, hyperbolic on one side and elliptic on the other. Furthermore, there are two kinds, depending on whether the $+A$ component has even or odd Milnor number.

The exceptional cases. We shall consider these only briefly (cf. [**1**]). The unfolding of the (E) germs can be given by

$$(E_m) - (ax + by + cxy + dy^2 + exy^2 + e_1y^3 + e_2xy^3), \quad m = 6, 8$$

$$E_7 - (ax + by + cx^2 + dxy + ey^2 + fx^2y).$$

In this expression we must hold $e_1 = e_2 = 0$ for $\pm E_6$. Figure 16 illustrates particular level curves and the resulting blossoms for E_6; Figure 17 does the same for E_7 and E_8.

We summarize these results in Table 3. It lists the simple catastrophes, representative blossoms, and all germs of one codimension lower which appear in their catastrophe sets. Following Arnold [**3**], we call the latter *bordering germs*.

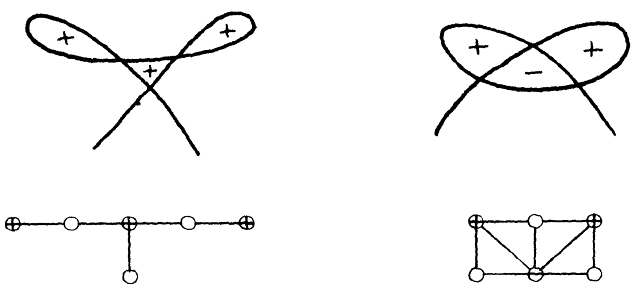
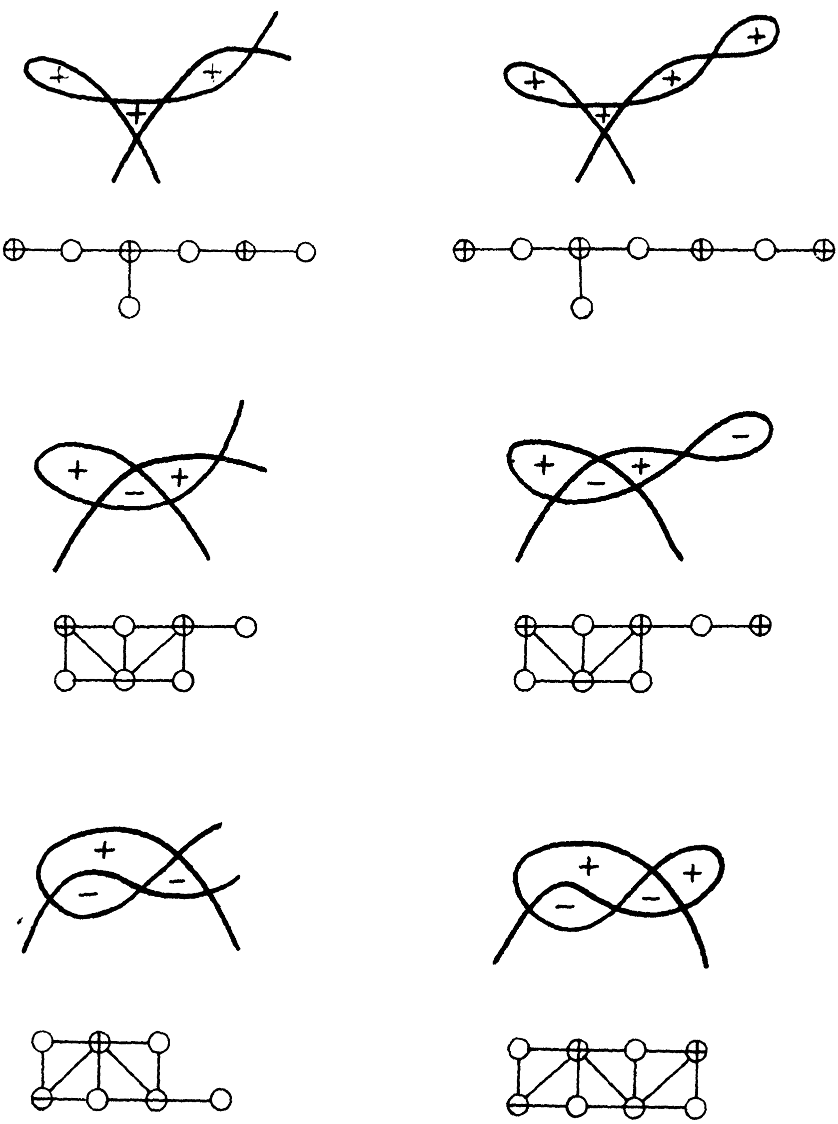


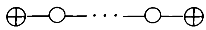
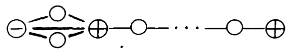
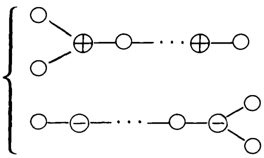
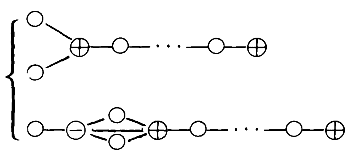
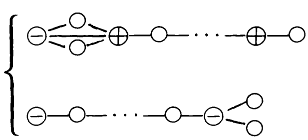
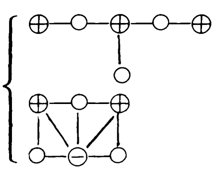
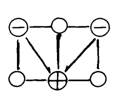
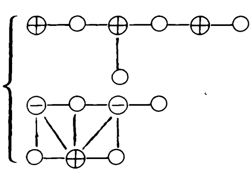
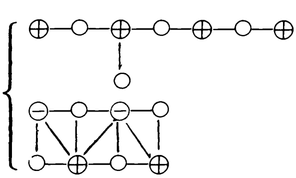
FIG. 16—E_6



E_7 E_8

FIG. 17

TABLE 3.

Catastrophe	Blossoms	Bordering germs
A_{2k}		A_{2k-1}, A^*_{2k-1}
A_{2k+1}		A_{2k}
A^*_{2k+1}		A_{2k}
D_{2k+2}		$\pm D_{2k+1}$, ($\pm A^*_3$ for D_4)
D^*_{2k+2}		$\pm D_{2k+1}$, A^*_{2k+1}, [$-A^*_{2k+1}$]
D_{2k+3}		D_{2k+2}, D^*_{2k+2}, A_{2k+2}
$-D_{2k+3}$		D_{2k+2}, D^*_{2k+2}, [$-A_{2k+2}$]
E_6		D_5, A_5, A^*_5
$-E_6$		$-D_5$, [$-A_5$, $-A^*_5$]
E_7		$\pm E_6$, D^*_6, A_6, [$-A_6$]
E_8		E_7, $\pm D_7$, A_7, [$-A_7$]

Suppose, following Arnold, we let $P \leftarrow Q$ denote the fact that P borders Q. Then Figure 18 puts the information in Table 3 in diagrammatic form. More important, it shows *all* types appearing in a

given catastrophe set. For instance, the catastrophe set of D_6 contains $\pm D_5$, D_4, D_4^*, A_4, A_3, A_3^*, A_2, and nothing else. In particular, D_6 does *not* contain the butterfly A_5^*, although D_6^* does. We shall return to this point later on. Figure 18 is the real version of Arnold's diagram [**2**, cor. 8.7; **4**, Fig. 1]; it is more complicated owing to the fact that a single complex function may have two inequivalent real forms (e.g., the distinction between minima and saddles does not apply in the complex case).

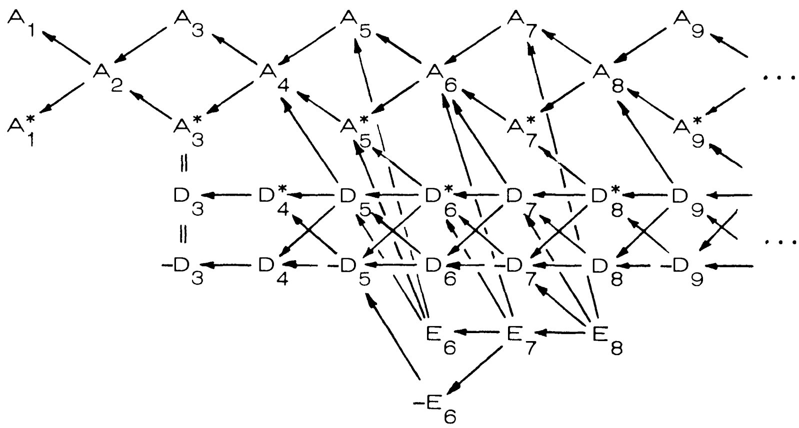


FIG. 18

5. Zeeman geometry.

The study of catastrophe germs $\chi : M \to \mathbb{R}^k$ is complicated by two facts: (1) M is non-linear; (2) except for the fold and the two cusps, the dimension k is too large to allow us to see the action of χ (e.g., Figures 5 and 6). If we merely want a picture of the catastrophe set $K \subset \mathbb{R}^k$, the list can be extended to the swallowtail and the first hyperbolic and elliptic umbilics. However, in applications and for developing geometric intuition it is helpful to have some device which makes it possible to visualize the higher-codimension catastrophes. For the simple germs we are considering, there is a convenient one: in a natural way, χ can be reduced to a family of plane maps $\alpha_c : \mathbb{R}^2 \to \mathbb{R}^2$.

To illustrate the procedure we shall consider first the corank 2 germs (these are the umbilics (D) and the exceptional (E) catastrophes); the cuspoids (A) require a slightly modified treatment. The universal unfoldings, as given in the last section, all have the form

$$F(a, b, c, x, y) = -ax - by + F^*(x, y, c).$$

Here, we are grouping the variables together as follows:

$$S = \mathbb{R}^2 : (x, y) \text{ state variables}$$

$$A = \mathbb{R}^2 : (a, b) \text{ first controls}$$

$$C = \mathbb{R}^{k-2} : c = (c_3, \ldots, c_k) \text{ other controls.}$$

The manifold M, defined as the locus $\partial F/\partial x = \partial F/\partial y = 0$, becomes the graph of the smooth map

$$T : S \times C \to A$$

$$T : \begin{cases} a = \dfrac{\partial F^*}{\partial x}(x, y, c) \\ b = \dfrac{\partial F^*}{\partial y}(x, y, c); \end{cases}$$

hence it is diffeomorphic to the source of T. Let $\theta : S \times C \to M$ denote this diffeomorphism. Define $\alpha : S \times C \to A \times C$ by the following diagram.

$$\begin{tikzcd} S \times C \arrow[r, "\theta"] \arrow[dr, "\alpha"'] & M \subset A \times C \times S \arrow[d, "\chi"] \\ & A \times C \end{tikzcd} \qquad \begin{tikzcd} (x, y, c) \arrow[r, mapsto, "\theta"] \arrow[dr, "\alpha"'] & \left(\dfrac{\partial F^*}{\partial x}, \dfrac{\partial F^*}{\partial y}, c, x, y\right) \arrow[d, "\chi"] \\ & \left(\dfrac{\partial F^*}{\partial x}, \dfrac{\partial F^*}{\partial y}, c\right) \end{tikzcd}$$

Suppose we let sing f denote the set of singular points of an arbitrary smooth map $f : N \to P$ of arbitrary manifolds N, P. Then, for the maps θ, χ, and α we are considering, it is not difficult to show that $\theta(\text{sing } \alpha) = \text{sing } \chi$ and $\alpha(\text{sing } \alpha) = \chi(\text{sing } \chi) = K$. See Fig. 19. Thus α, with the advantage of a linear source, can replace χ in studying the geometry of K.

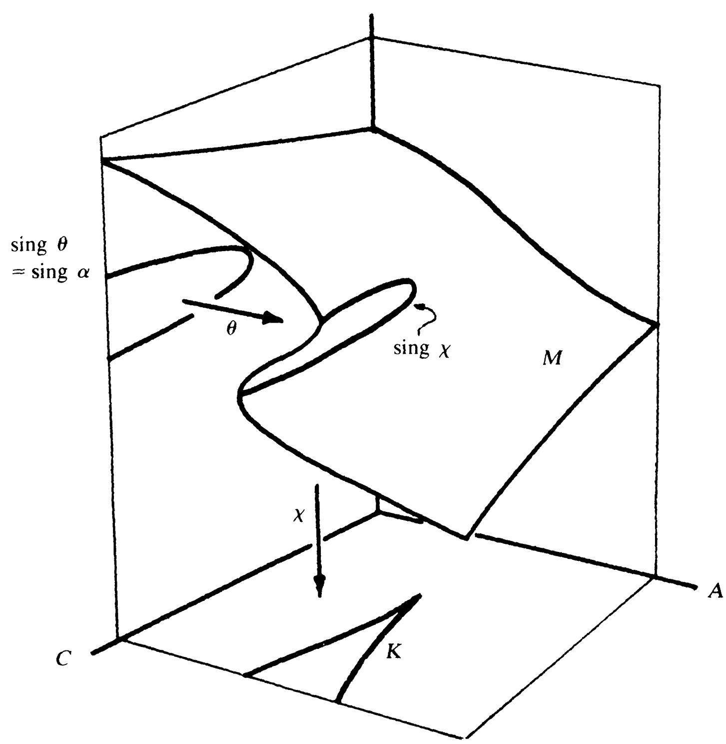


FIG. 19

Finally, the particular form of $\alpha : S \times C \to A \times C$ permits it to be interpreted as a $(k-2)$ parameter family

$$\alpha_c : S \to A \qquad\qquad c \in C = \mathbb{R}^{k-2}$$

$$(x, y) \to \left(\frac{\partial F^*}{\partial x}(x, y, c), \frac{\partial F^*}{\partial y}(x, y, c)\right).$$

These are the plane maps we shall study. As Figure 20 suggests, θ, χ, and M can be fibered in the same way, giving

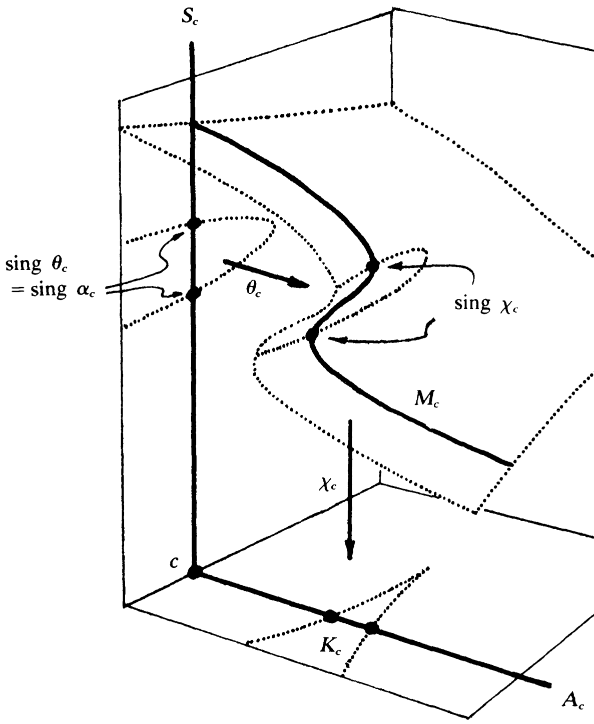


FIG. 20

$$\theta_c\,(\text{sing}\,\alpha_c) = \text{sing}\,\chi_c \subset M_c$$

$$\alpha_c\,(\text{sing}\,\alpha_c) = \chi_c\,(\text{sing}\,\chi_c)$$

$$K = \bigcup_c \alpha_c\,(\text{sing}\,\alpha_c) = \bigcup_c \chi_c\,(\text{sing}\,\chi_c).$$

That is, the apparent contour of each α_c in $A = \mathbb{R}^2$ is a one-dimensional section of K. We want to know how these sections vary with the parameter c.

To adapt the foregoing to the cuspoid $(A)_{k+1}$, unfold it by

$$F(a, b, c, x, y) = \pm \frac{x^{k+2}}{k+2} + y^2 \mp \left(ax + \frac{bx^2}{2} + \frac{c_3x^3}{3} + \cdots + \frac{c_kx^k}{k}\right)$$

and then set

$$S = \mathbb{R}^2 : (x, b)$$

$$A = \mathbb{R}^2 : (a, b)$$

$$C = \mathbb{R}^{k-2} : c = (c_3, \ldots, c_k).$$

By this trick, M can once again be identified with a graph

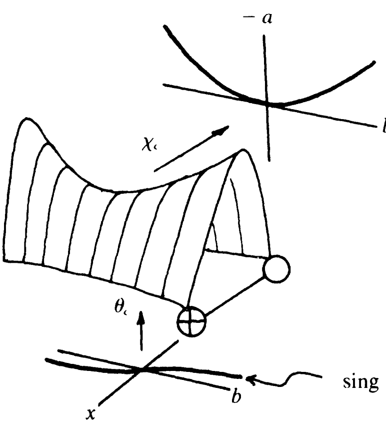
$$T : S \times C \to A$$

$$T : \begin{cases} a = x^{k+1} - (bx + c_3x^2 + \cdots + c_kx^{k-1}) \\ b = b. \end{cases}$$

The rest follows as before. We illustrate all this with several examples.

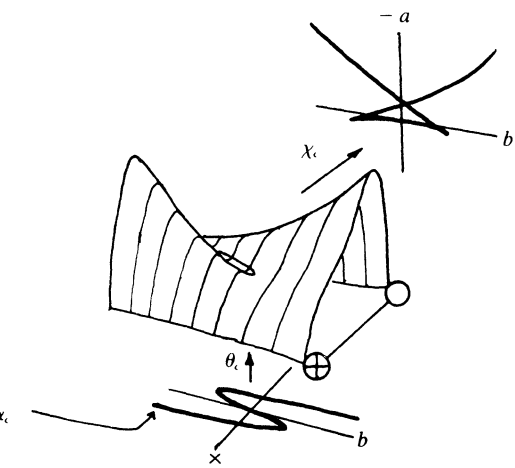
The swallowtail A_4. Here there is only one other control variable $c_3 = c$, leading to a one-parameter family

$$\alpha_c : \begin{cases} a = x^4 - bx - cx^2 \\ b = b. \end{cases}$$



α_c ; $c < 0$

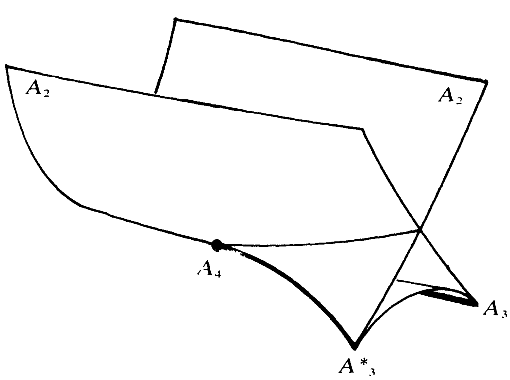
FIG. 21



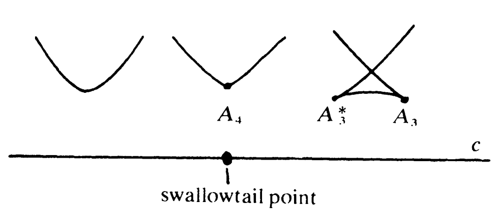
α_c ; $c > 0$

FIG. 22

The singular set sing α_c is the curve $b = 4x^3 - 2cx$ and consists entirely of fold points when $c < 0$ (Fig. 21). If $c > 0$, there are two cusps on the fold line, of type A_3 at $x = +\sqrt{c/6}$ and type A_3^* at $x = -\sqrt{c/6}$ (Fig. 22; cf. [**6**, Fig. 15]). These figures demonstrate that α_c factors through θ_c and χ_c, and that the sections M_c, as well as the full manifold M, consist of blossoms of an unfolding.



(a)



(b)

FIG. 23

Figure 23a shows these slices sandwiched together to give the complete catastrophe set K of A_4. In particular it illustrates nicely just how the germs A_2, A_3, A_3^* border A_4. (Incidentally, this surface was already studied extensively in the last century; see Klein [**10**, page 99].) The simpler tableau (Fig. 23b) apparently conveys less information. However, it shall ultimately be more useful, because we can adapt it to the higher catastrophes, where no complete picture of K is possible. Such a tableau has been used to sketch the four-dimensional geometry of the parabolic umbilic [**16**, page 87].

The butterfly A_5. In this case,

$$F(a,b,c,d,x,y)=\frac{x^6}{6}+y^2-\left(ax+\frac{bx^2}{2}+\frac{cx^3}{3}+\frac{dx^4}{4}\right)$$

and

$$\alpha_{c,d}:\mathbb{R}^2\to\mathbb{R}^2$$
$$a=x^5-bx-cx^2-dx^3$$
$$b=b.$$

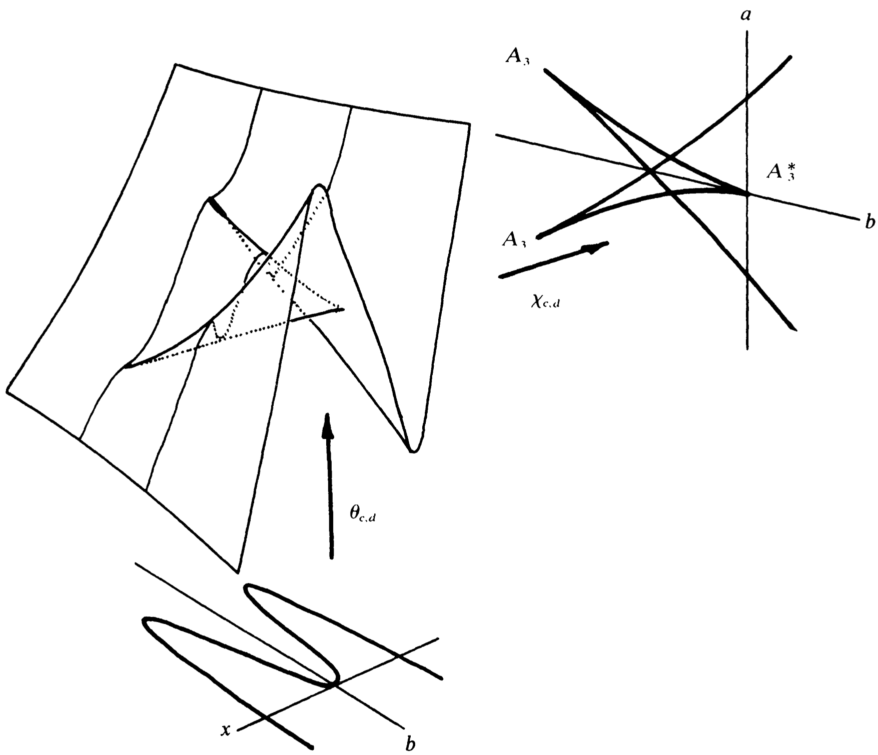


FIG. 24

The particular section which gives A_5 its name is shown in Figure 24; it occurs when $c=0$, $d>0$. As the controls vary, either of the A_3 points can move toward the A_3^* point, first piercing the A_2 curve and eventually fusing with A_3^* in a swallowtail. Exactly how this happens is displayed in the tableau of Figure 25. The dotted curves bordering on A_5 are not new catastrophes. They represent non-generic sections of its catastrophe set (viz, those where a cusp point meets a fold line), and must be brought into the tableau to separate inequivalent generic sections, like (1) and (3). See Fig. 26. Besides cusp-fold intersections, other non-generic sections include the lip and the beak-to-beak singularities to be discussed later, with the umbilics. All these cases are an artifact of the method of sections we are using and must be expected in all the tableaus.

The tableaus of A_5 and its dual A_5^* are identical, except that cusps and dual cusps must be interchanged in the sections. Thus Figure 24 would have one A_3 and two A_3^* cusps, and the central lozenge would contain only two minima, instead of three, as it does for A_5. Applications of the butterfly, and further discussion of its geometry, are given in [**21**, **22**].

A_6: Notice that although the unfolding

$$F(a, b, c, d, e, x, y) = \frac{x^7}{7} + y^2 - \left(ax + \frac{bx^2}{2} + \frac{cx^3}{3} + \frac{dx^4}{4} + \frac{ex^5}{5}\right)$$

is five-dimensional in this case, we can still construct a tableau. It appears in Figure 27. The swallowtail

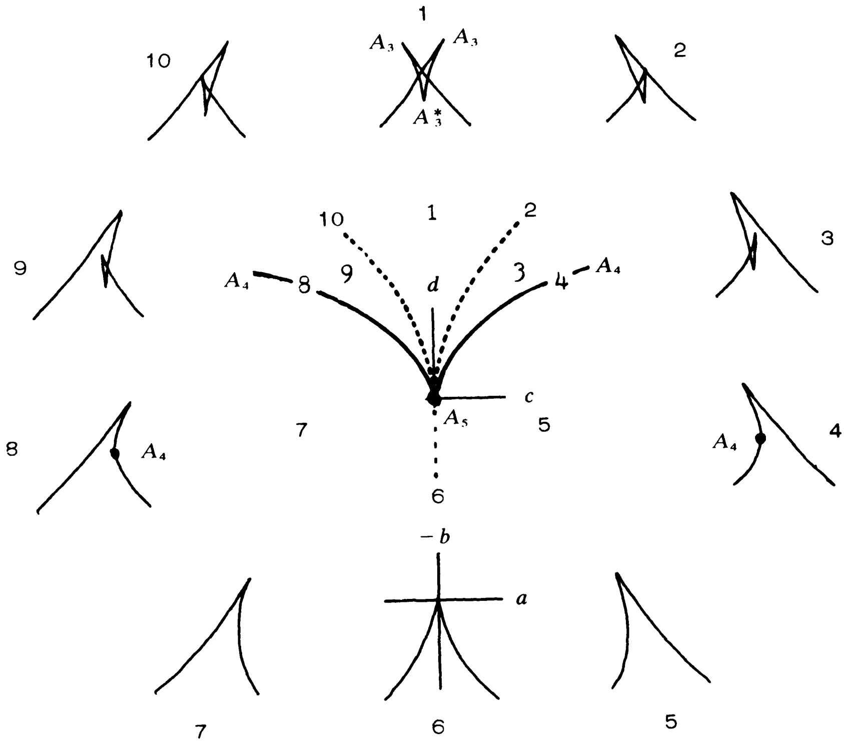


FIG. 25

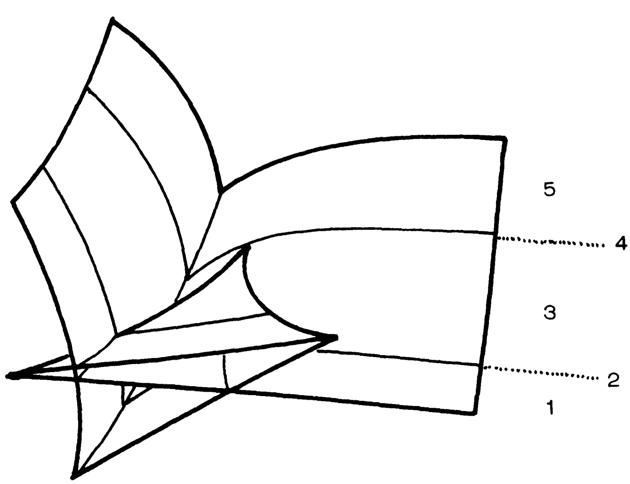


FIG. 26

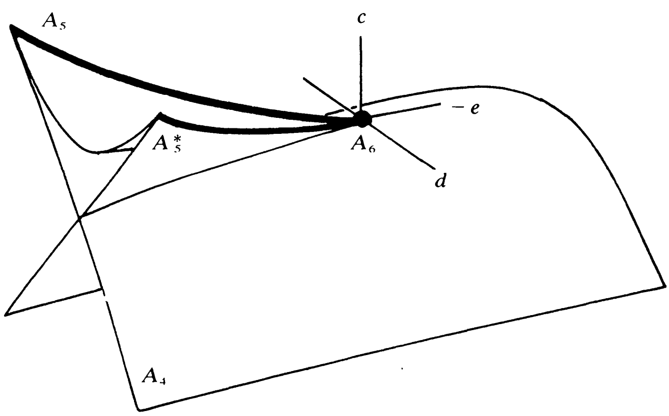


FIG. 27

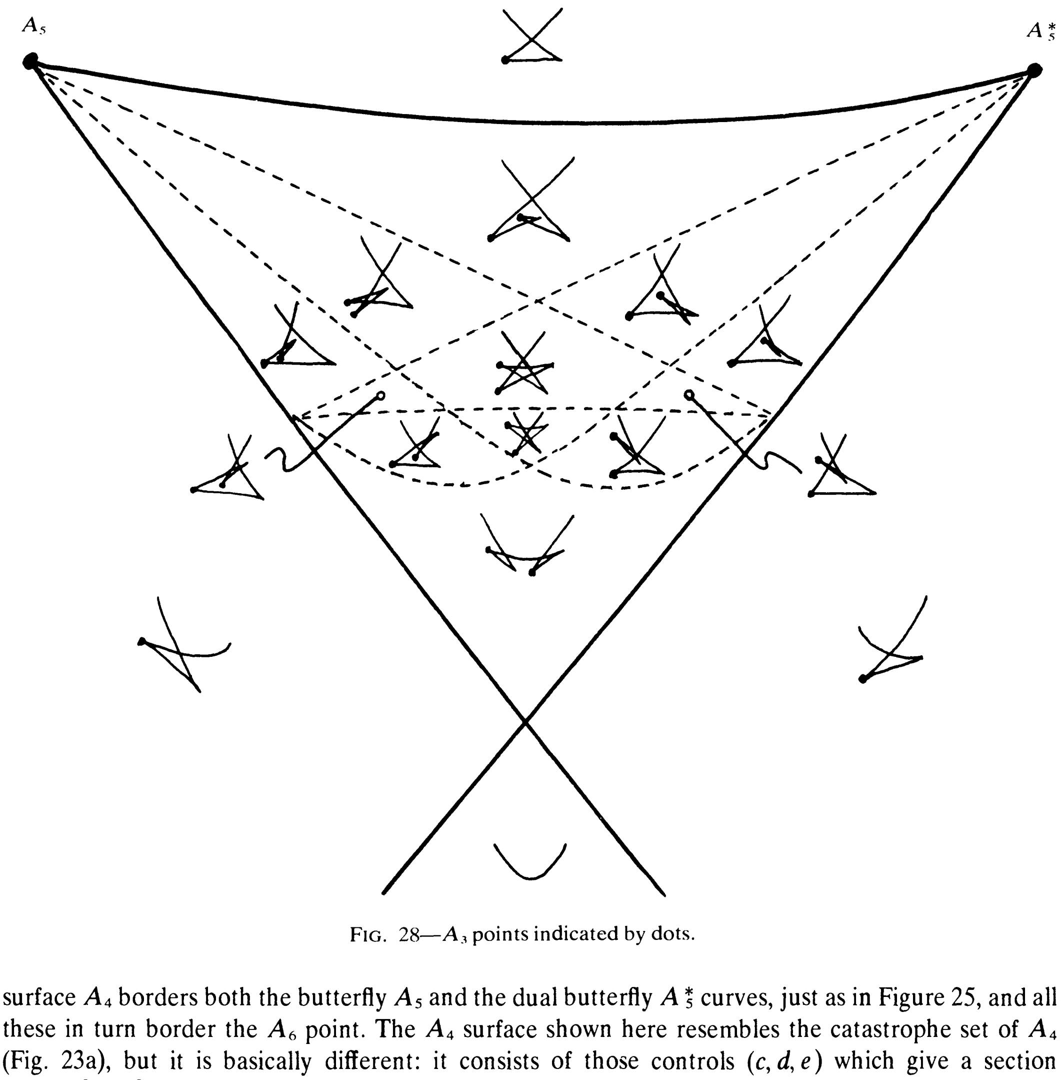


FIG. 28—A_3 points indicated by dots.

surface A_4 borders both the butterfly A_5 and the dual butterfly A_5^* curves, just as in Figure 25, and all these in turn border the A_6 point. The A_4 surface shown here resembles the catastrophe set of A_4 (Fig. 23a), but it is basically different: it consists of those controls (c, d, e) which give a section $\alpha_{c,d,e} : \mathbb{R}^2 \to \mathbb{R}^2$ having a swallowtail point.

A tableau containing all sections of A_6 is too unwieldy. Figure 28, which pictures all the generic sections, is adequate, as the rest can be interpolated visually. The dotted lines represent surfaces of non-generic sections bordering on A_6. Compare neighborhoods of A_5 and A_5^* in Figure 28 with Figure 25.

There are at least two ways of extending the tableau method. First, the two-dimensional sections can be assembled into three-dimensional ones (as in Fig. 26). One such section ($e>0$, d near 0) for A_6 is given in Figure 29; it essentially becomes part of a new two-dimensional tableau of three-dimensional sections.

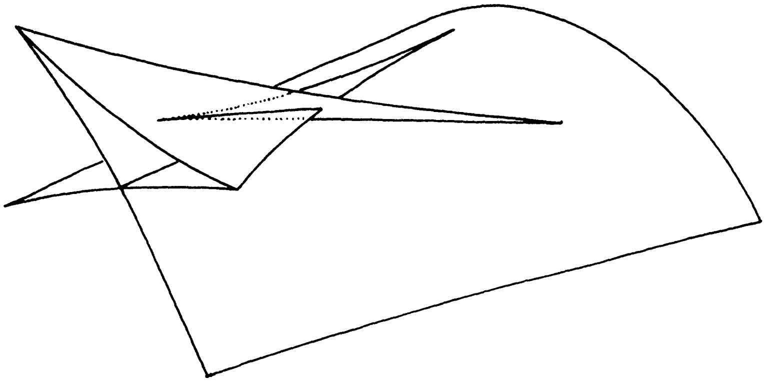


FIG. 29

Second, it is possible to picture the entire A-series in an inductive sequence of tableaus, by splitting up the space of other controls into a product of planes—i.e., writing $c=((c_3,c_4),(c_5,c_6),\ldots)$.

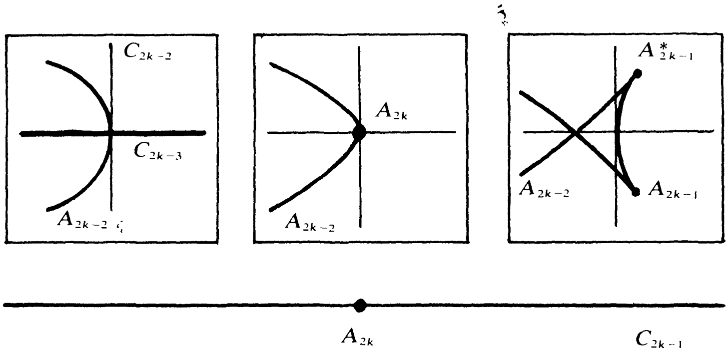


FIG. 30

Figure 30 does this for the odd-codimension germs A_{2k}, which perhaps ought to be called the "swallowtailoids." It depicts some of the underlying geometry of the catastrophe set K of A_{2k} and in particular it explicates Arnold's bordering diagram

$$\begin{tikzcd} & A_{2k-1} \arrow[dl] & \\ A_{2k-2} & & A_{2k} \arrow[ul] \arrow[dl] \\ & A^*_{2k-1} \arrow[ul] & \end{tikzcd}$$

Figure 31 does the same thing for the even germs A_{2k+1}, and suggests that the term "cuspoid" refers more properly to them alone. Arnold's bordering diagram is

$$\begin{tikzcd}[column sep=small] A_{2k-1} & & A_{2k+1} \arrow[dl] \\ & A_{2k} \arrow[ul] \arrow[dl] & \\ A^*_{2k-1} & & \end{tikzcd}$$

Interchanging $A_{2k\pm1}$ and $A^*_{2k\pm1}$ in Figure 31 gives the geometry of the dual cuspoid A^*_{2k+1}. Note the inductive tableau in each of these figures.

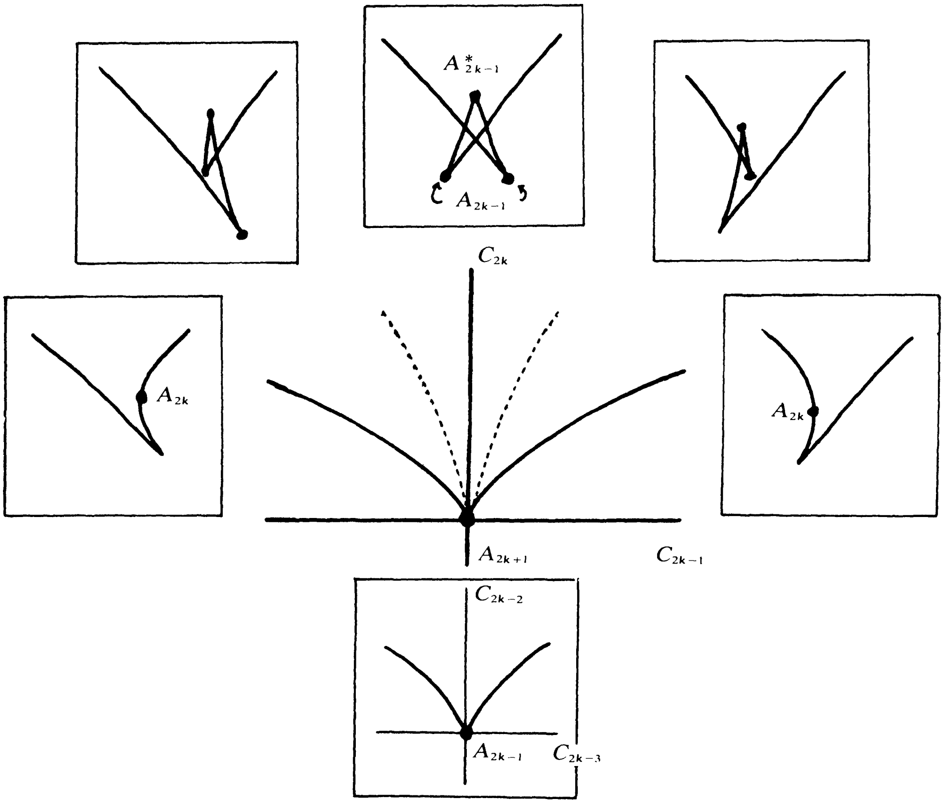


FIG. 31

6. Sketching umbilics.

We have analyzed the catastrophe map $\chi : M \to \mathbb{R}^k$ by introducing a diffeomorphism $\theta : \mathbb{R}^k \to M$ and taking sections of the composite $\alpha = \chi \circ \theta$, as in the following diagram.

$$\begin{tikzcd} \mathbb{R}^2 \arrow[r, "\theta_c"] \arrow[dr, "\alpha_c"'] & M_c \subset \mathbb{R}^4 \arrow[d, "\chi_c"] \arrow[d, shift left=6ex, "\pi"] & (x, y, a, b) \arrow[d, "\pi"] \\ & \mathbb{R}^2 = \mathbb{R}^2 & (a, b) \end{tikzcd}$$

In fact the sequence $\mathbb{R}^2 \to M_c \to \mathbb{R}^2$ cannot be visualized directly, as M_c lies in $\mathbb{R}^4$. We must first project $\mathbb{R}^4$ to $\mathbb{R}^3$ along the y-axis.

$$\begin{tikzcd}
 & M_c \arrow[d, "j"] & \subset \mathbb{R}^4 \arrow[d, "\pi_y"] & (x, y, a, b) \arrow[d, "\pi_y"] \\
\mathbb{R}^2 \arrow[ur, "\theta_c"] \arrow[r, "\hat{\theta}_c"'] \arrow[dr, "\alpha_c"'] & \hat{M}_c \arrow[d, "\hat{\chi}_c"] & \subset \mathbb{R}^3 \arrow[d, "\pi_x"] & (x, a, b) \arrow[d, "\pi_x"] \\
 & \mathbb{R}^2 & = \mathbb{R}^2 & (a, b)
\end{tikzcd}$$

As we have already seen, the restriction $j: M_c \to \hat{M}_c$ is an isometric embedding for the A-germs (and we simply identified M_c with $\hat{M}_c$). The pictures of the last section were thus obtained from $\mathbb{R}^2 \to \hat{M}_c \to \mathbb{R}^2$ with no loss of information. However, the blossoms of the other catastrophes involve the second variable y in an essential way (this is what corank 2 means); in consequence, the map $j: M_c \to \hat{M}_c$ can have singularities. But for the umbilics at least, $\hat{M}_c$ has a particularly simple form, and the resulting sequence

$$\alpha_c : \mathbb{R}^2 \xrightarrow{\hat{\theta}_c} \hat{M}_c \xrightarrow{\hat{\chi}_c} \mathbb{R}^2$$

is easy to visualize. Our aim is to describe $\hat{M}$ and obtain sections of the catastrophe sets K of the D-germs, as apparent contours of $\mathbb{R}^2 \to \hat{M}_c \to \mathbb{R}^2$. (This approach is implicit in Figures 16–25 of [**6**].)

Let us begin with the "progenitor" of the umbilics, the (non-determinate) germ $D_\infty = xy^2$. (The Jacobian ideal contains no power of the maximal ideal, so the germ has infinite codimension. See [**5, 8, 18**] for discussions of determinacy.) The partial unfolding

$$F(a, b, x, y) = xy^2 - ax - by$$

leads to the map $\hat{\theta}: \mathbb{R}^2 \to \mathbb{R}^3$

$$\hat{\theta}: \begin{aligned} a &= y^2 \\ b &= 2xy \\ x &= x, \end{aligned}$$

see Fig. 32. Whitney [**19**] has shown that $\hat{\theta}$ represents the only stable singularity $\mathbb{R}^2 \to \mathbb{R}^3$; its image $\hat{M}$ is sometimes called the **Whitney umbrella** (although this name is usually applied to the algebraic set

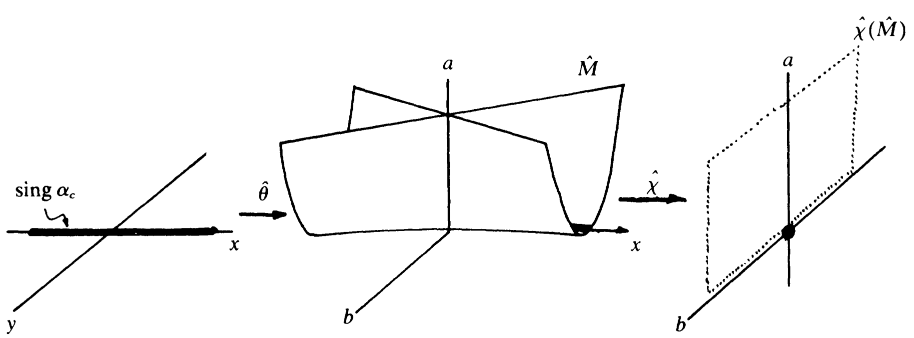


FIG. 32

$b^2 = 4ax^2$, which is the union of $\hat{M}$ and the a-axis). The projection $\hat{\chi}$ carries $\hat{M}$ two-to-one onto the upper half-plane, except that the entire x-axis is squashed down to the origin. The composite $\alpha = \hat{\chi} \circ \hat{\theta}$ is extremely unstable; sing α is the x-axis and $\alpha(\text{sing } \alpha)$ is the origin.

The road to stability for D_∞ begins by making it determinate: it becomes $(D)_{m+2}$ when we add $\pm x^{m+1}/(m+1)$ to the germ and the unfolding. The new immersion $\hat{\theta}: \mathbb{R}^2 \to \mathbb{R}^3$ is

$$\hat{\theta}: \begin{aligned} a &= \pm x^m + y^2 \\ b &= 2xy \\ x &= x. \end{aligned}$$

To get its image, just deform the Whitney umbrella by bending the x-axis into the curve $a = \pm x^m$. This can even be done for $m = 1$; what results is the germ D_3, equivalent to A_3^*. Figure 33 illustrates the four main possibilities. Notice that the parabolic type looks hyperbolic on one side, elliptic on the other—just like its blossoms. Note also the apparent contours of these new umbrellas.

The last step is to bring in the full universal unfolding

$$F(a, b, c_3, \ldots, c_{m+1}, x, y) = \pm\left(\frac{x^{m+1}}{m+1} - \frac{c_4x^2}{2} - \cdots - \frac{c_{m+1}x^{m-1}}{m-1}\right) + xy^2 - ax - by - c_3y^2$$

$$\hat{\theta}_c: \begin{aligned} a &= \pm x^m + y^2 \mp (c_4x + \cdots + c_{m+1}x^{m-2}) \\ b &= 2y(x - c_3) \\ x &= x \end{aligned}$$

of $(D)_{m+2}$, and see how the other controls affect $\hat{M}_c$. While the details will take some time to unravel, the basic facts are straightforward. Two features of $\hat{M}_c$ can be altered: the position of its **double-line** of self-intersection, and the shape of its **keel**—i.e., the curve $a = \pm x^m$. As we shall see, c_3 moves the double line and the remaining controls unfold the keel.

This approach provides us with an overview of the entire D-series: the various sections $\alpha_c : \mathbb{R}^2 \to \mathbb{R}^2$ describe all possible ways of stabilizing the projection of the Whitney umbrella given in Figure 32.

The hyperbolic umbilic D_4. The keel $a = x^2$ is already stable, so there is only one other control $c_3 = c$. Figure 34 illustrates its effect on the double line and on the apparent contour. The surface $\hat{M}_c$, which like M_c is a configuration of blossoms, has been marked to show the regions corresponding to minima, maxima, and saddles. A boundary between two regions consists of degenerate critical points, arising from minimum-saddle (A_2) and maximum-saddle ($-A_2$) collisions. These boundaries are the singular points of $\hat{\chi}_c : \hat{M}_c \to \mathbb{R}^2$ and hence their images make up the apparent contour of $\hat{\chi}_c$. See Figure 35, and compare it with Figure 5.16 in [**16**, page 77].

Of course the maximum-saddle ($-A_2$) collision is not catastrophic, and must be ignored in constructing sections of the catastrophe set K in the ab-plane. The apparent contours of all $\hat{\chi}_c$ are sandwiched together in Figure 36; (a) shows the full contour and (b) shows the $-A$ points removed to give the strict catastrophe set of D_4.

The situation here is typical of all the umbilics; that is, the singular points of $\hat{\chi}_c$ which lie in back of the double line (i.e., for which $x < 0$) are non-catastrophic. In the future we shall indicate them with dotted lines or omit them altogether. This will simplify the diagrams considerably.

The elliptic umbilic D_4^*. Again there is but a single other control $c_3 = c$; its effect is illustrated in Figure 37 (which is a more straightforward version of [**6**, Fig. 20]). The catastrophe set of D_4^* and the evolution of its blossom are shown in Figure 38.

The parabolic umbilics D_5 and $-D_5$. A maximum of D_5 is a minimum of $-D_5$, so the discarded components of the apparent contour of $\hat{\chi}_c : \hat{M}_c \to \mathbb{R}^2$ (i.e., the ones shown as dotted lines) make up the catastrophe set of $-D_5$. It will therefore be convenient to look at the entire contour, keeping track of which parts belong to D_5 and which to $-D_5$.

The keel $a = x^3$ is now unstable; it is unfolded by the second control $c_4 = d : a = x^3 - dx$. (Incidentally, the "flabbiness" of the flabby pleat [**6**, section 6] is due entirely to the instability of this keel. Compare D_3 and D_{2k+3} in Figure 33.) We can see immediately just how the elliptic and hyperbolic umbilics border on $\pm D_5$: set $d > 0$ and set c so that the double line of $\hat{M}_{c,d}$ appears at the

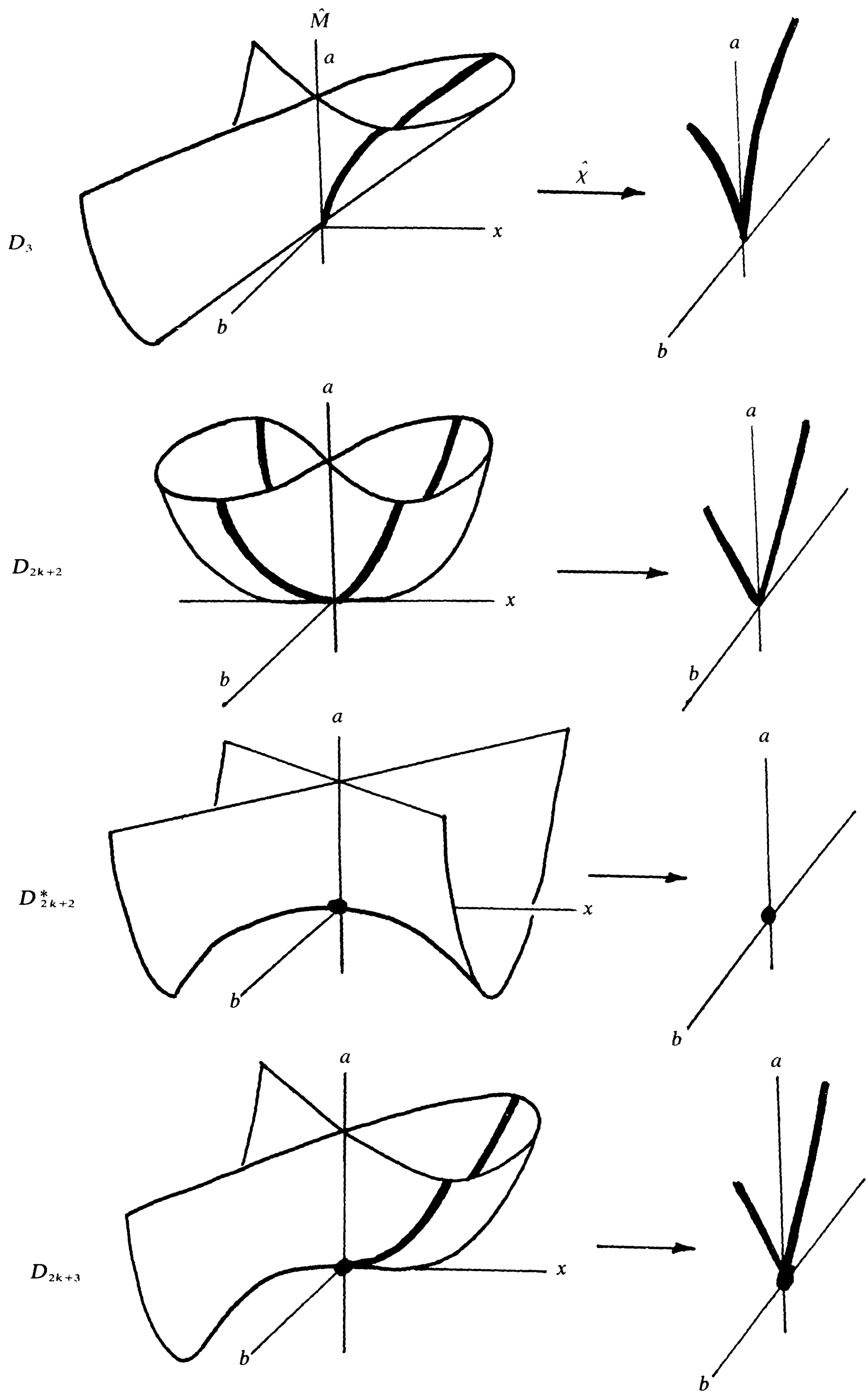


FIG. 33

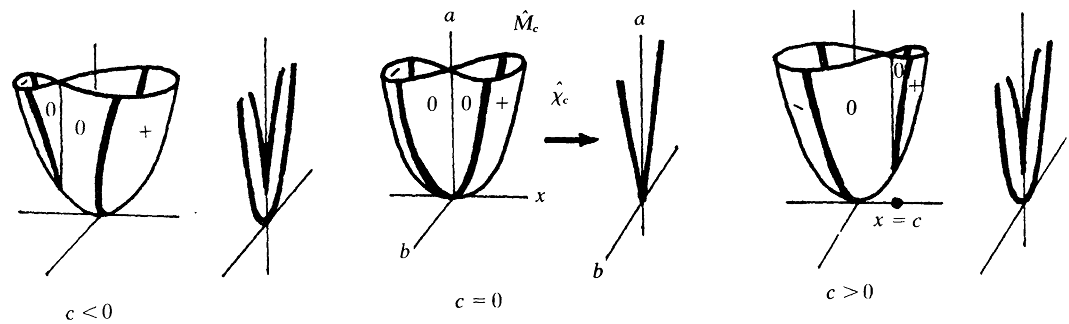


FIG. 34

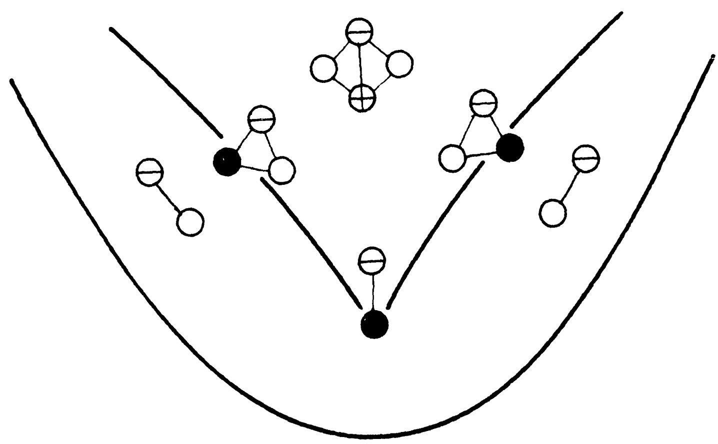


FIG. 35

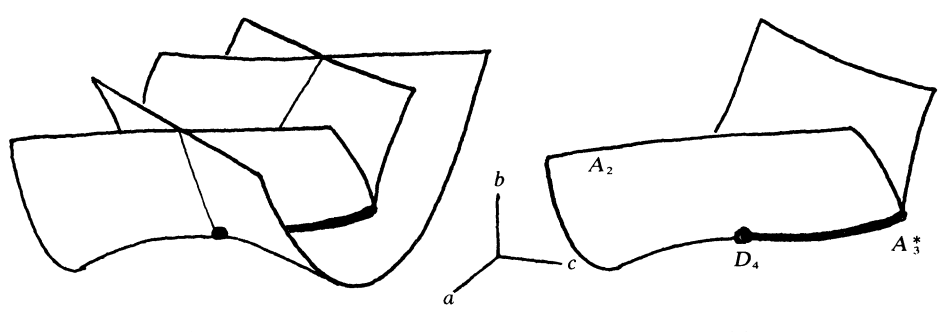


FIG. 36

maximum or the minimum of $a = x^3 - dx$ (Fig. 39). Since the critical points of $a = x^3 - dx$ occur when $3x^2 - d = 0$, the D_4 and D_4^* points of K are $d = 3c^2$, $c > 0$ and $c < 0$, respectively.

The **lip** and **beak-to-beak** (***b–b***) singularities also occur for the first time. They are due to the instability of the keel, i.e., they are given by $d = 0$. Figure 40 illustrates the beak-to-beak, and Figure 41 the lip. The former appears in D_5 and the latter in $-D_5$.

The other catastrophe bordering on D_5 is the swallowtail A_4. It does not border on $-D_5$; this is

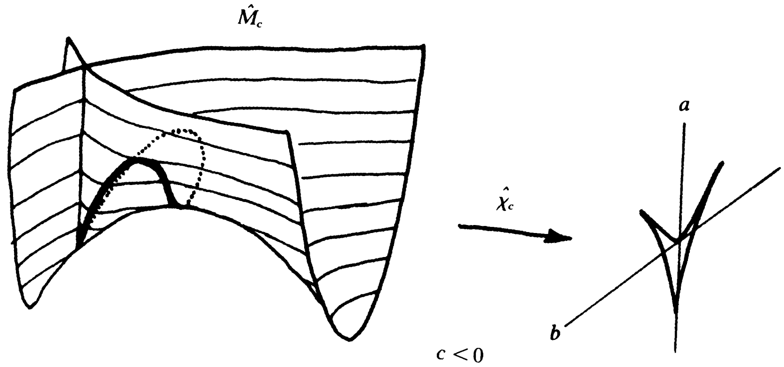


FIG. 37

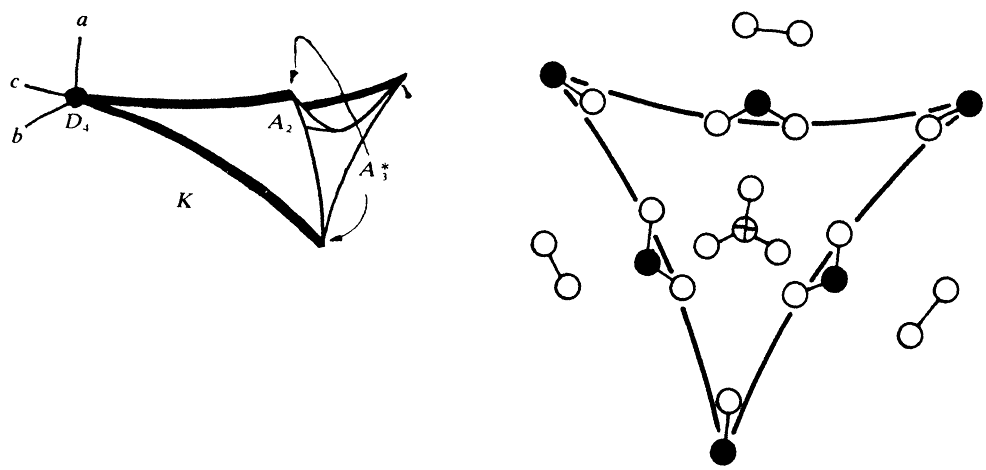


FIG. 38

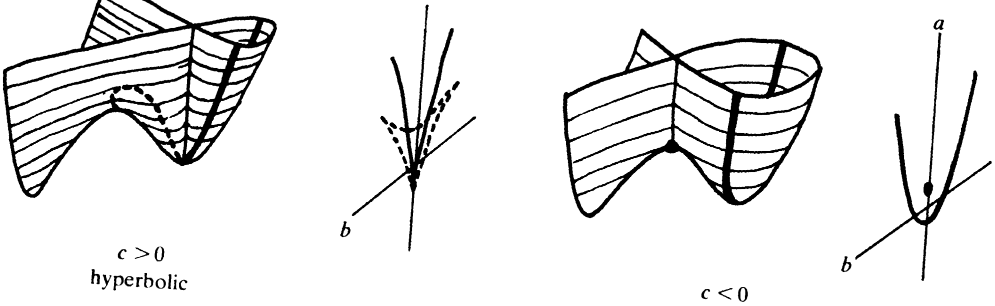
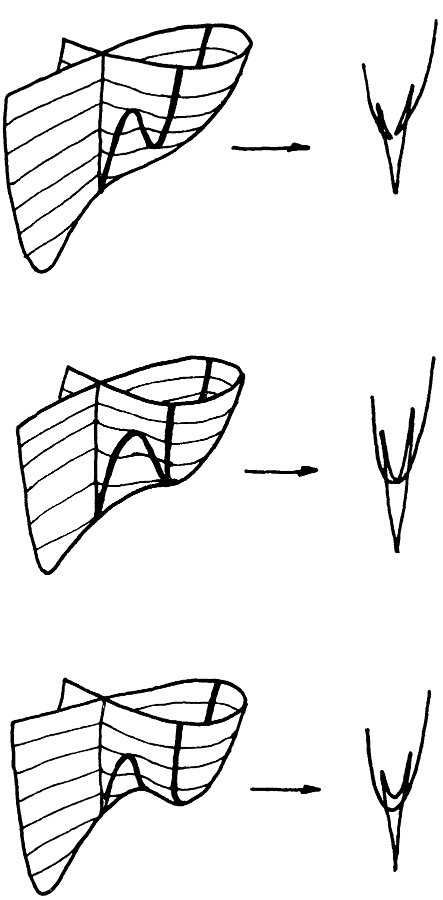
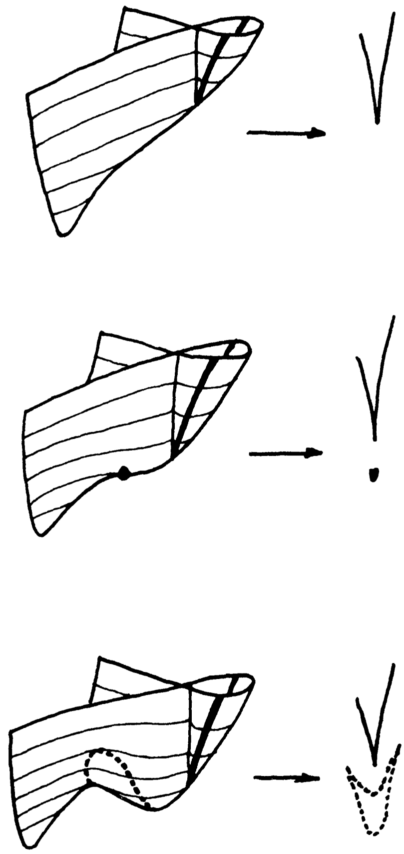


FIG. 39



beak-to-beak

FIG. 40



lip

FIG. 41

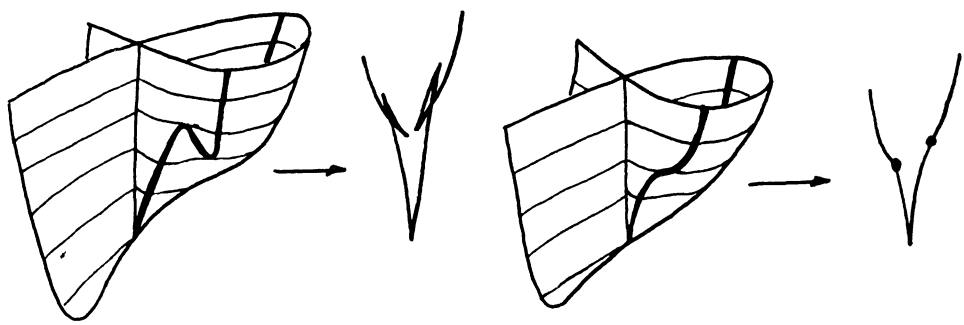


FIG. 42

asserted by Arnold's diagram and it can also be deduced from the shape of $\hat{M}$. It occurs when the singular set, on $\hat{M}$, has a horizontal inflection. Since

$$a = \frac{5}{2}x^3 - \frac{3}{2}cx^2 - \frac{3}{2}dx + \frac{1}{2}cd$$

on the singular set, the swallowtail condition is $d = -c^2/5$, $c < 0$ (Fig. 42).

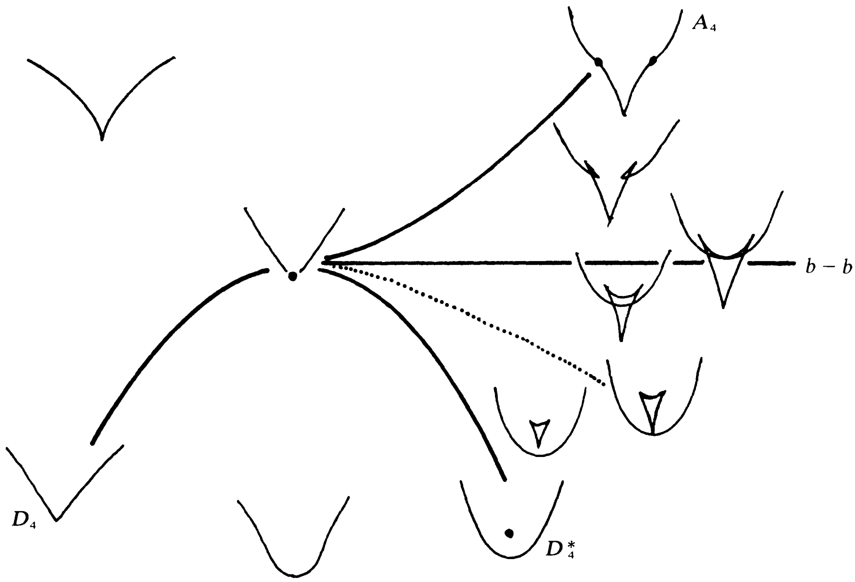
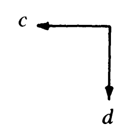


Tableau for D_5



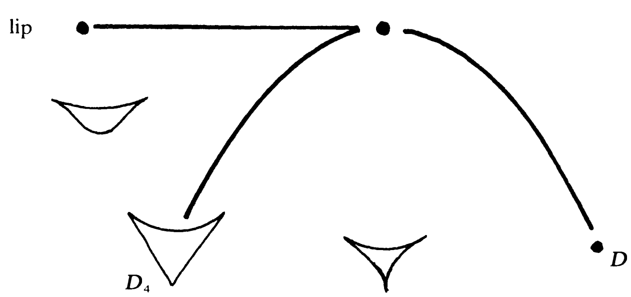


Tableau for $-D_5$

FIG. 43

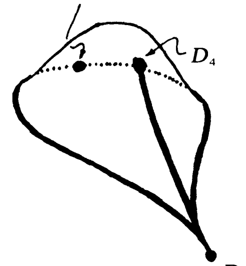


FIG. 44

Tableaus for D_5 and $-D_5$ are given in Figure 43. Notice that there is one non-generic piercing section in D_5 but none in $-D_5$. Of course, if these tableaus are combined, as in Thom's book, additional piercings arise when components of D_5 and $-D_5$ meet non-generically. By referring to $\hat{M}$ it can be seen that the side vertices, but not the bottom vertex, of the $-D_5$ section can pierce the D_5 section. On the other hand, only the bottom vertex of the compact (triangular) component of D_5 can pierce its non-compact component. These details about the shape of K come naturally out of its representation as an apparent contour.

A sandwich of $-D_5$ sections is given in Figure 44; the complete catastrophe set is just a (non-linear) cone in $\mathbb{R}^4$ on this mandolin-shaped figure.

The hyperbolic umbilic D_6. The unfolding we use is

$$F(a, b, c, d, e, x, y) = \frac{x^5}{5} + xy^2 - \left(ax + by + cy^2 + \frac{dx^2}{2} + \frac{ex^3}{3}\right);$$

$$\begin{aligned} a &= x^4 + y^2 - dx - ex^2 \\ \hat{\theta} : b &= 2y(x - c) \\ x &= x. \end{aligned}$$

As before, c controls the double line of $\hat{M}$ and (d, e) unfolds the keel: $a = x^4 - dx - ex^2$.

We can quickly get a rough idea how the catastrophe set sits in the tableau. First, as in the case of $\pm D_5$, an umbilic occurs when the double line of $\hat{M}_{c,d,e}$ is positioned over a critical point of the keel. In other words, (c, d, e) gives an umbilic if c is a critical point of $x^4 - dx - ex^2$ — i.e., a solution of $4c^3 - d - 2ec = 0$. But this is just the equation defining the catastrophe surface M of the cusp A_3 (Fig. 6). Next, the umbilic is parabolic, or there is a lip or a beak-to-beak singularity, if the keel has an x^3-type point. This happens precisely when keel controls (d, e) lie on the familiar cusp. See Fig. 45;

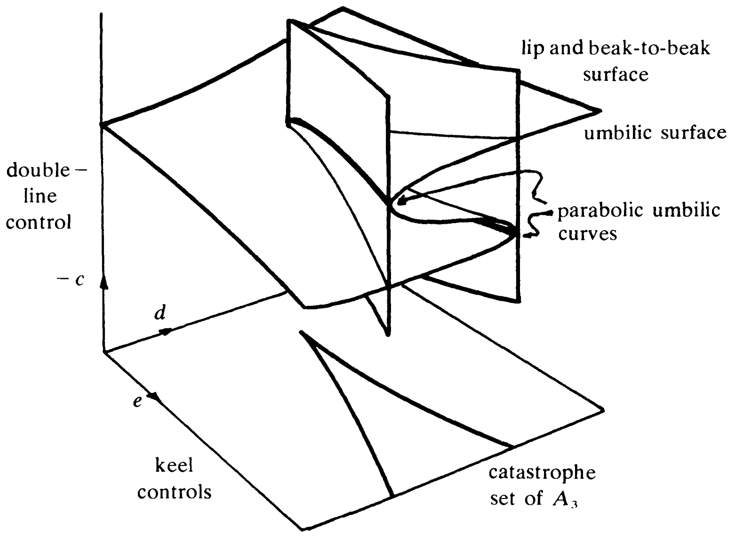


FIG. 45

since $(D)_{m+2} = x \,.\, (A)_{m-1}$ it should not be particularly surprising to find the catastrophe manifold of A_3 in the tableau for D_6. In general, the catastrophe manifold of A_{m-1} is the locus of umbilics in the tableau for D_{m+2}; (cf. Fig. 43).

As it happens, the larger, outer sheet of the umbilic surface is hyperbolic, and the smaller middle section is elliptic. The left-hand curve of parabolic umbilics is type $-D_5$; the sheet lying above it is a

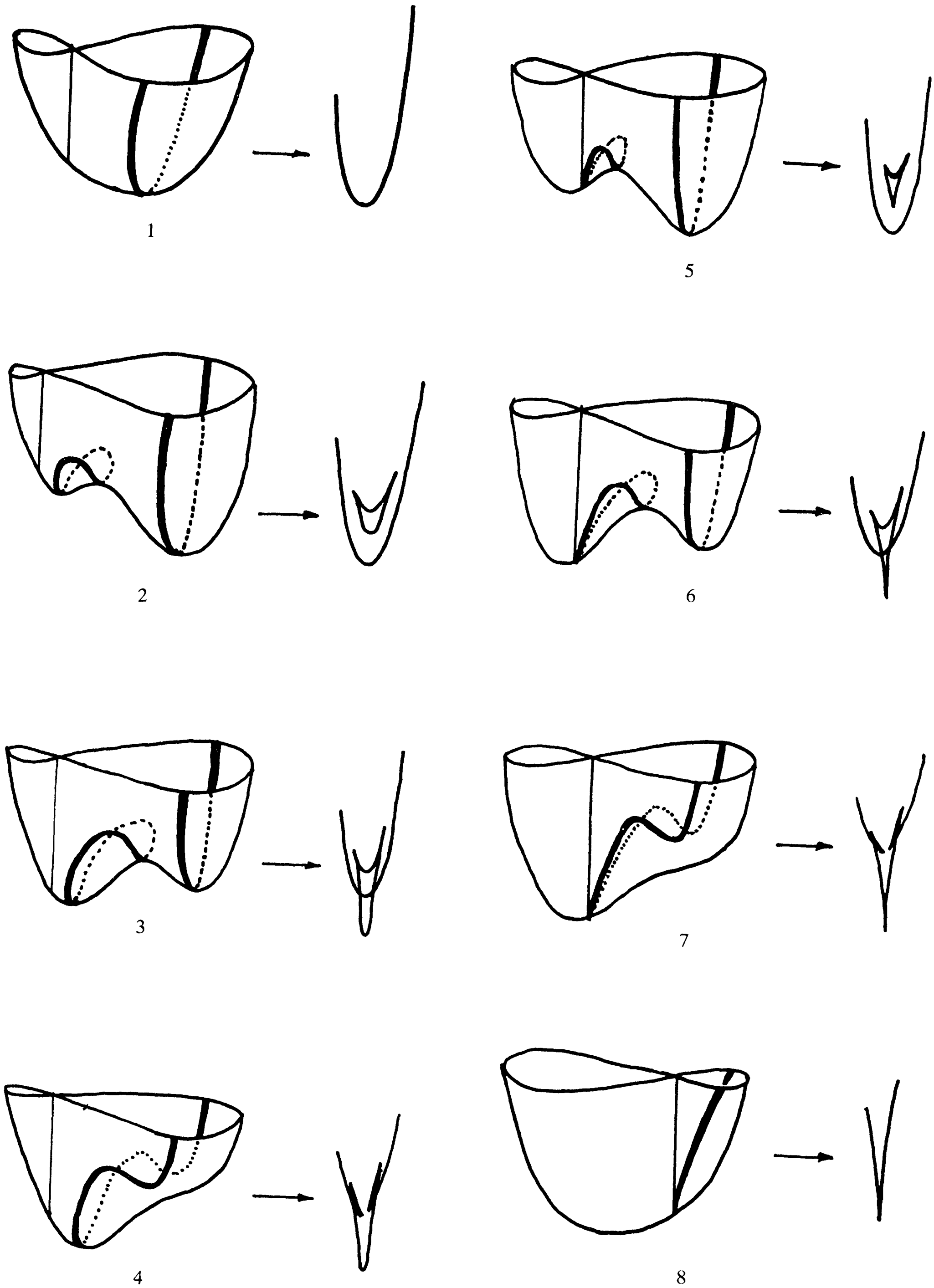


FIG. 47

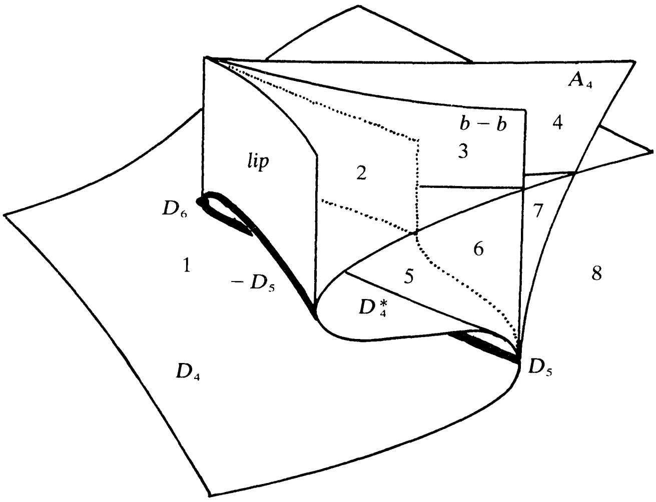


FIG. 46—Tableau for D_6

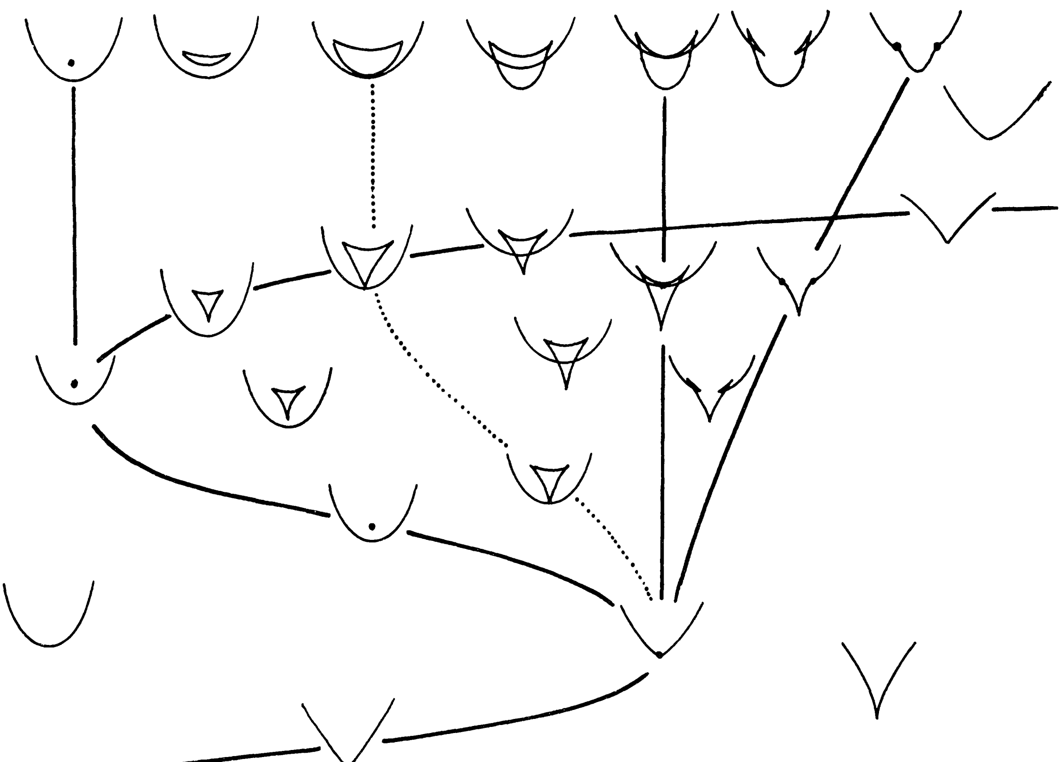


FIG. 48

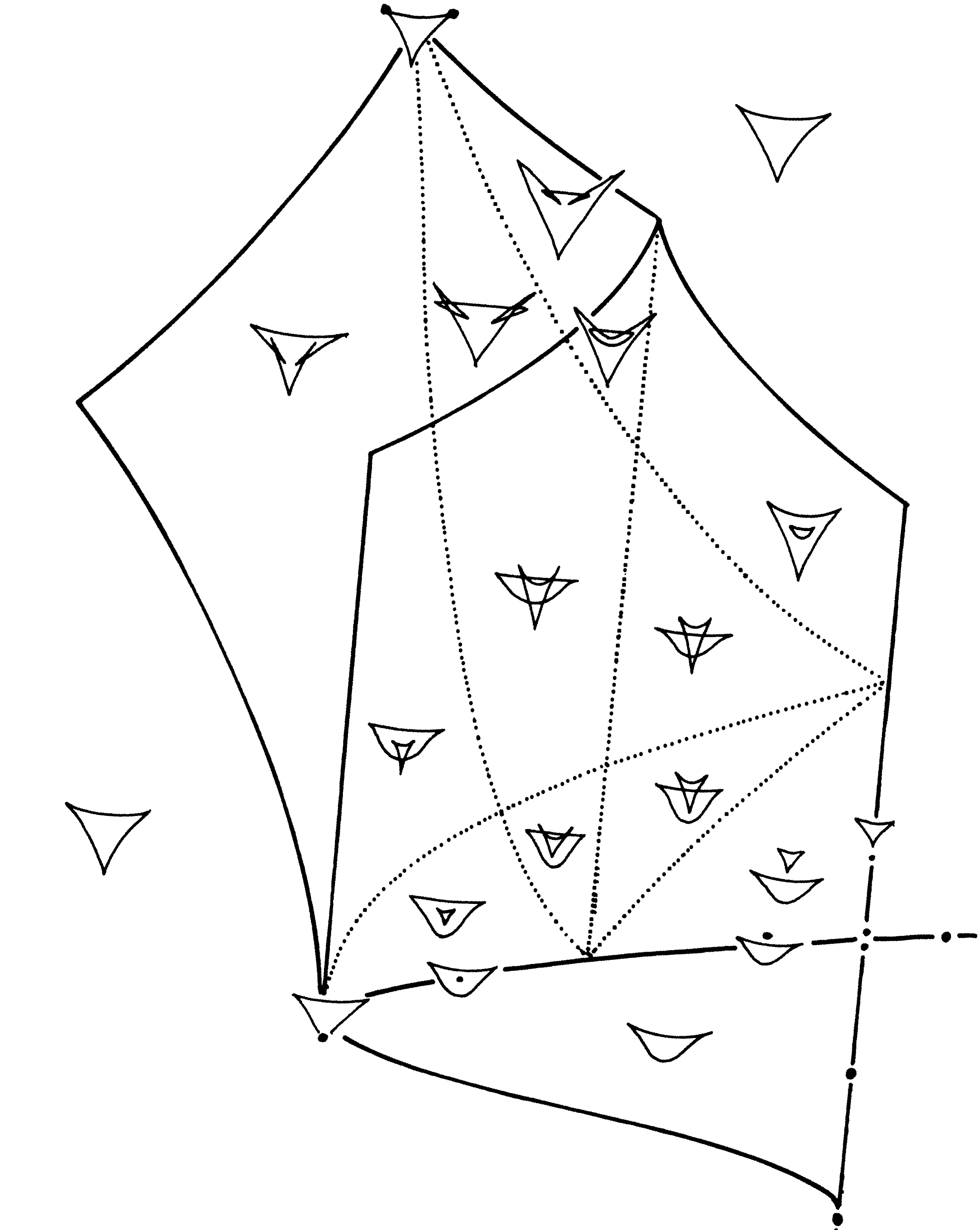


Fig. 50

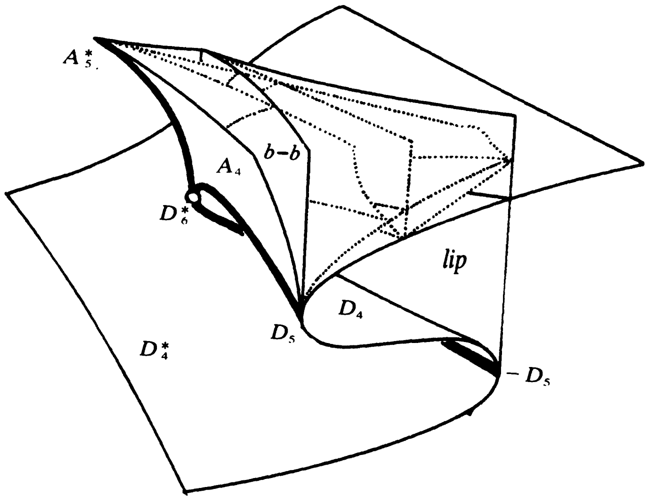


FIG. 49—Tableau for D_6^*

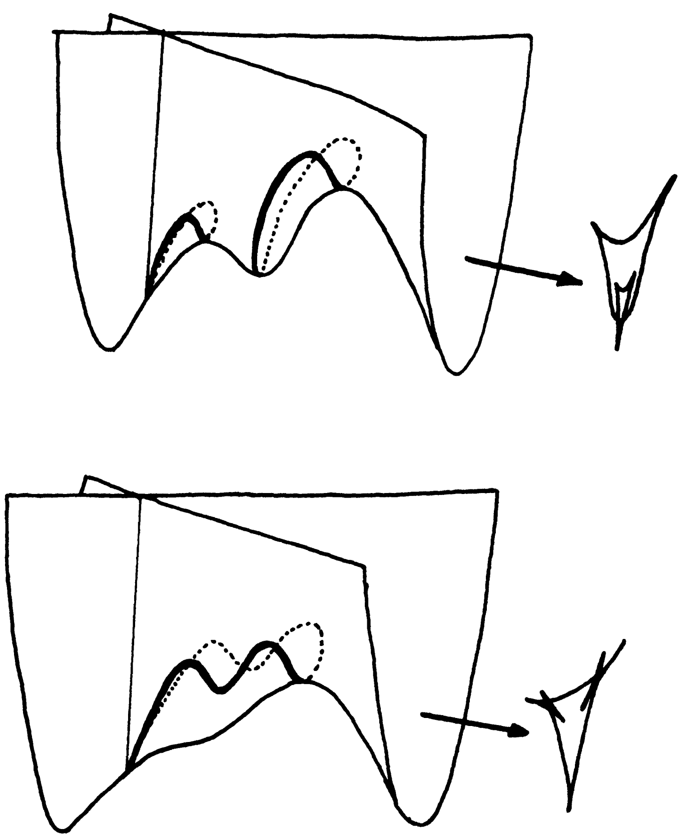


FIG. 51

lip, but the one below is a non-catastrophic beak-to-beak and must be deleted. The other parabolic umbilic curve is D_5; the sheet above is a beak-to-beak, the one below a non-catastrophic lip. There is also an A_4 swallowtail surface, which terminates along the D_5 curve and the line of intersection of the lip and beak-to-beak surfaces. Finally, there is a surface of piercing sections. This more detailed picture of the D_6 tableau appears in Figure 46. It contains altogether eight disjoint open sets; the map $\hat{\chi}_{c,d,e} : \hat{M}_{c,d,e} \to \mathbb{R}^2$ for each is given in Figure 47. Lastly, all inequivalent sections of D_6 are displayed in Figure 48. Notice that the butterflies A_5 and A_5^* do not appear, even though they have smaller codimension than D_6. This agrees with Arnold's bordering diagram. Notice also that the configurations around the D_5 and $-D_5$ points agree with Figure 43.

The elliptic umbilic D_6^*. Taking into account the change

$$a = -x^4 + y^2 + dx + ex^2,$$

there is much overlap with D_6. In particular, all the inferences contributing to Figure 45 continue to hold here. However, in sorting through the details, we find that the D_4 and D_4^*, D_5 and $-D_5$, and lip and beak-to-beak regions must be interchanged. The swallowtail surface follows D_5 to the other side, and in doing so it develops a cusp edge consisting of A_5^* dual butterfly points. There are also several piercing surfaces to consider. See Figure 49. Sections are displayed over the end and top of the tableau in Figure 50. A sample of maps $\hat{\chi}_c : \hat{M}_c \to \mathbb{R}^2$ is given in Figure 51. Notice how the change in the shape of the keel, from D_6 to D_6^*, opens the way for A_5^* points to appear.

It is possible to construct an inductive chain of tableaus for the D-series, similar to the A-series. Also, tableaus of the other catastrophes can be constructed. Figure 52 is one for E_6.

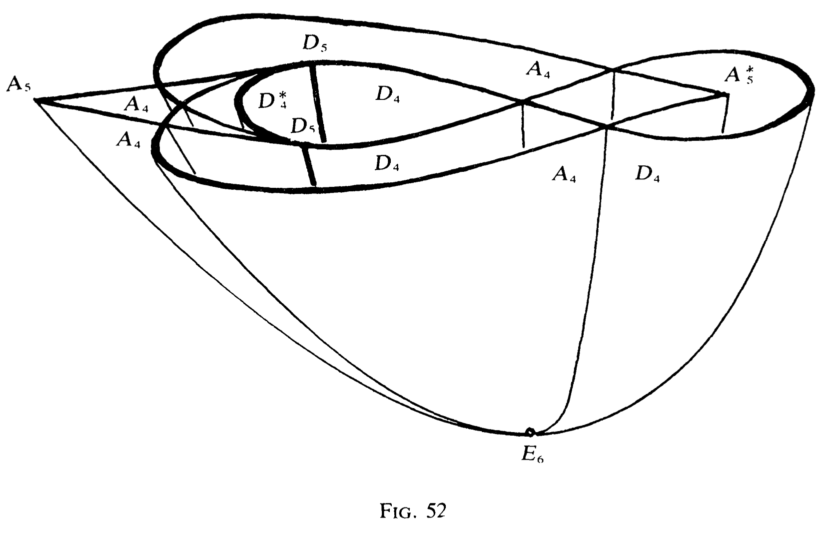


FIG. 52

This work was supported by FPC Grant 3782.

References

1. N. A'Campo, Le groupe de monodromie du déploiement des singularités isolées de courbes planes I, Math. Ann., 213 (1975) 1–32.

2. V. I. Arnold, Normal forms for functions near degenerate critical points, the Weyl groups for A_k, D_k, E_k and Lagrangian singularities, Func. Anal. Appl., 6 (1973) 245–272.

3. ———, Remarks on the stationary phase method and Coxeter numbers, Russian Math. Surveys, 28 (1973) 19–48.
4. ———, Critical points of smooth functions, Proc. Int. Cong. Math. Vancouver, 1974, 19–39.
5. T. Brocker, Differentiable Germs and Catastrophes, (translated by L. Lander) Camb. Univ. Press, 1975.
6. J. Callahan, Singularities and plane maps, this MONTHLY, 81 (1974) 211–240.
7. D. Gromoll and W. Meyer, On differentiable functions with isolated critical points, Topology, 8 (1969) 361–370.
8. M. Golubitsky and V. Guillemin, Stable Mappings and their Singularities, Graduate Test in Mathematics, Springer, New York, 1974.
9. S. M. Gusein–Zade, Dynkin diagrams for singularities of functions of two variables, Func. Anal. Appl., 8 (1974) 295–300.
10. F. Klein, Elementary Mathematics from an Advanced Standpoint: Arithmetic, Algebra, Analysis (translated by E. R. Hedrick & C. A. Noble), Dover, New York, 1939.
11. J. Mather, Right equivalence, Warwick University preprint, 1969.
12. ———, Stability of C^∞ mappings III: Finitely determined map-germs, Publ. Math. I.H.E.S., 35 (1968) 127–156.
13. T. Poston and A. E. R. Woodcock, A higher catastrophe machine, Math. Proc. Camb. Phil. Soc. 79 (Pt 2) 343–350.
14. D. Siersma, The singularities of C^∞-functions of right-codimension smaller than or equal to eight, Indag. Math., 35 (1973) 31–37.
15. I. N. Stewart, The geometry of binary quartic forms (part I), Warwick University preprint, 1976.
16. R. Thom, Structural Stability and Morphogenesis, (translated by D. H. Fowler), Benjamin, New York, 1975.
17. R. Thom and E. C. Zeeman, Catastrophe theory: its present state and future perspectives, in Dynamical Systems, Warwick 1974; Springer Lecture Notes in Mathematics 468, 336–401.
18. D. J. A. Trotman and E. C. Zeeman, The classification of elementary catastrophes of codimension $\leqq 5$, in Symposium on Catastrophe Theory Notes (Seattle 1975), Springer Lecture Notes in Mathematics, 525.
19. H. Whitney, On singularities of mappings of Euclidean spaces, in Symposium International de Topologia Algebraica, Mexico City, 1958.
20. A. E. R. Woodcock and T. Poston, A Geometrical Study of the Elementary Catastrophes, Springer Lecture Notes in Mathematics 373, 1974.
21. E. C. Zeeman, Applications of catastrophe theory, in Intl. Conference on Manifolds, Tokyo Univ. 1973.
22. ———, Catastrophe theory, Sci. Amer., (April 1976) 65–83.
23. ———, A catastrophe machine, Towards a Theoretical Biology 4, ed. C. H. Waddington, Edinburgh Univ. Press 1972.
24. ———, The umbilic bracelet and the double-cusp catastrophe, Symposium on Catastrophe Theory (Seattle 1975), Springer Lecture Notes in Mathematics, 525.

DEPARTMENT OF MATHEMATICS, SMITH COLLEGE, NORTHAMPTON, MA 01063.

MATHEMATICAL NOTES

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Department of Mathematics, University of Wisconsin, Madison, WI 53706.

COMMENTS AND COMPLEMENTS

RICHARD A. BRUALDI

The purpose of this article is to provide periodically an opportunity for our readers to comment on articles that have been published in the Mathematical and Classroom Notes sections of the MONTHLY.

Because of the nature of the material we seek to publish, it is not surprising that some duplication of results and ideas already in the literature occurs. Such duplication may not be undesirable, depending upon availability, presentation, and so on. In this regard, we are pleased to present the following information received from our readers during the past year.

Calculus. In the article *A property of the logarithm* (this MONTHLY, 72(1965) 767; reprinted in *Selected Papers on Calculus*, Math. Assoc. of America (1969) 121–122) D. S. Greenstein gives a simple demonstration that $\lim_{x\to\infty}(\ln x/x)=0$ using only the inequality (1) $0<\ln x<x\,(x>1)$. H. Anton has shown us the following argument for $\lim_{x\to\infty}x/e^x=0$ which can be presented early in elementary calculus: Since $\ln x=2\ln x^{1/2}$, it follows from (1) that $0<\ln x<2\sqrt{x}$, thus $1<x<e^{2x^{1/2}}$, therefore $1/e^x<x/e^x<e^{2x^{1/2}-x}$. Since $\lim_{x\to\infty}(2\sqrt{x}-x)=-\infty$ (by rationalizing), the limit is readily deduced.

Analysis. In two recent articles in Volume 83, 1976, of the MONTHLY—S. K. Stein, *An inequality in two monotone functions*, 469–471, and I. S. Rusza, *Probabilistic generalization of a number-theoretical inequality*, 723–725—an old inequality of Tchebyshev has been rediscovered. For correspondence on this matter we are grateful to H. Falk, P. S. Bullen, K. Jogdeo, and S. Stein who has summarized for us the information he has obtained concerning it. The inequality in question is the following: let f, g be monotonic (in the same sense) functions and let v and h be non-negative measurable functions. Then

$$\int_a^b f(v(t))g(v(t))h(t)dt\cdot\int_a^b h(t)dt \geqq \int_a^b f(v(t))h(t)dt\cdot\int_a^b g(v(t))h(t)dt.$$

An alternate formulation in terms of covariance is $\mathrm{Cov}\,[f(x),g(x)]\geqq 0$, where x is a real-valued random variable. As Jogdeo pointed out, it follows by induction and conditioning that if h and k are real functions on R^n, nondecreasing in each of the n arguments separately and if $x_1,\dots,x_n$ are independent random variables, then $\mathrm{Cov}\,[f(x_1,\dots,x_n),g(x_1,\dots,x_n)]\geqq 0$. A list of references include: (1) T. Apostol, *Mathematical Analysis*, 2nd edition, Addison-Wesley, Reading, 1974; (2) J. D. Esary, F. Proschan, and D. W. Walkup, *Association random variables with applications*, Annals Math. Stat., 38(1967) 1466–1474; (3) C. M. Fortuin, P. W. Katelyn, and J. Gimbre, *Correlation inequalities on some partially ordered sets*, Commun. Math. Phys., 22(1971) 89–103; (4) G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge, 1952; (5) C. H. Kimberling, *Some corollaries to an integral inequality*, this MONTHLY, 81(1974) 269–270. In (5) the inequality is used to obtain other inequalities, including results on power series. In (3) the inequality is developed for distributive lattices and applied to Ising ferromagnetics in a magnetic field. A proof, depending on the fact that the integral of a non-negative function is non-negative, is outlined in Exercise 7.17, p. 177 of (1). Theorem 43, p. 43 of (4) contains a discrete version. In addition, R. O. Davies has communicated to Stein the following proof that "time-average of (velocity)^2" is not larger than "distance-average of (velocity)^2":

$$\begin{aligned}\text{Time-average of (velocity)}^2 &\leqq (\text{Distance-average of velocity})^2\\ &\leqq \text{Distance-average of (velocity)}^2,\end{aligned}$$

each of the two inequalities being a special case of Schwarz's inequality.

There is a review written by R. B. Burckel (in Math. Reviews, 52(1976) 1200–1201, No. 8433) of D. J. Newman's *An entire function bounded in every direction* (this MONTHLY, 83(1976) 192–193) which contains some information which we would like to pass on to our readers. Newman constructs a non-constant entire function $f(z)$ which, in spite of Liouville's theorem, is bounded on every line through 0. Burckel points out that the function $F(z)=f(z)f(iz)$ is a non-zero entire function such that $F(z)\to 0$ as $z\to\infty$ along any line through 0. In addition he points out that other reasonably elementary accounts of functions having the property of $f(z)$ above can be found in *Aufgaben und Lehrsätze aus der Analysis* by G. Pólya and G. Szegö (4th edition vol. I (1970), vol. II (1971)) and *Infinitesimal Calculus* (English translation) by J. Dieudonné (Paris (1971), p. 282, Exer. 20). According to Burckel, these latter constructions go back to G. Mittag-Leffler, J. Malmquist, E. LeRoy, and E. Lindelöf and have the property of F above, besides exhibiting other pathology.

Concerning the article *The Green's function and determining equations* (this MONTHLY, 82(1975) 747–749) by D. Sanchez, I. Stakgold has observed that equation (1) of that article does not hold for all solutions as claimed. Sanchez has replied that that equation holds only for those solutions having average value equal to zero.

Concerning his article with M. Simonovits entitled *On a problem of Hirschhorn* (this MONTHLY, 83(1976) 23–26) P. Erdös writes that T. Šalát has pointed out a minor error in the calculations. The multiplicative factor $r_n = q_{n+1} - q_n$ is missing from the remainder terms in formulae (7) and (8). Therefore the part from formula (7) to formula (9) should be changed as follows:

$$r_{n+1} = r_n + (r_n/q_n) + O(r_n^2/q_n^2). \tag{7}$$

Further by (7)

$$\begin{aligned} d_{n+1} - d_n &= (s_{n+1} - s_n) - (q_{n+1} - q_n) = n(r_{n+1} - r_n) + r_{n+1} - r_n \\ &= (n+1)(r_{n+1} - r_n) = \left(1 + \frac{1}{n}\right)\frac{s_n}{q_n} + O\left(\frac{r_n s_n}{q_n^2}\right). \end{aligned} \tag{8}$$

(A) First we need that r_n/q_n tends to zero. A trivial induction gives that $q_n > n$. Therefore $r_n = o(q_n)$ if r_n is bounded. If r_n is not bounded, then it tends to infinity, since it is (trivially) monotone increasing. Hence $q_n/n \to \infty$ and thus by (4) $r_n = o(n)$; hence in any case $r_n = o(q_n)$.

Thus from (8)

$$r_{n+1} - r_n = (1 + o(1))\frac{r_n}{q_n} = (1 + o(1))\frac{s_n}{nq_n} \tag{9}$$

and from here on the proof of (2), (3) and their sharper form in (C) and (D) are unchanged.

In his article *Another elementary proof of Peano's existence theorem* (this MONTHLY, 83(1976) 556–560), C. Gardner gave a proof of Peano's theorem without using integrals, uniform continuity or uniform convergence. MONTHLY readers may be interested also in the article *Elementary proofs of Peano's existence theorem* (J. Australian Math. Soc., 15(1973) 366–372) by M. A. Dow and R. Výborný, where a proof of Peano's theorem is given, also not using uniform continuity or uniform convergence.

C. L. Belna has written that the theorem proved by L. M. Levine in *On a necessary and sufficient condition for Riemann integrability* (this MONTHLY, 84(1977) 205) follows from some old results of W. H. Young. The interested reader can consult Young's articles *On some applications of semi-continuous functions* (Atti del IV Congresso Internazionale dei Matematici (Roma 1908) II, Roma, 1909, pp. 49–60) and *On the discontinuities of a function of one or more real variables* (Proc. London Math. Soc. (2), 9(1909) 117–124).

Algebra. The short proof of the fundamental theorem of algebra given by F. Terkelsen (this MONTHLY, 83(1976) 647) is similar to proofs already in the literature. A. M. Gendler has given us as reference *Principles of Mathematical Analysis* by W. Rudin (2nd edition, McGraw-Hill, 1964, p. 170). W. C. Waterhouse has contributed the references *Modern Algebra and Matrix Theory* by O. Schreier and E. Sperner (Chelsea, 1959, p. 238) and *Linear Algebra* by S. Lang (Addison-Wesley, 1969, p. 374). In addition Waterhouse writes that a proof of this nature was given by Argand around 1806 (see S. S. Petrova, *Sur l'histoire des démonstrations analytiques du théorème fondamental de l'algèbre*, Historia Math., 1(1974) 255–261).

We are grateful to Daniel Mark Rosenblum, a graduate student at Carnegie-Mellon University, for observing some connections among some recent articles in the MONTHLY. In *On the commutativity of rings* (this MONTHLY, 71(1964) 267–271) R. Ayoub and C. Ayoub prove the following. Let R be a ring and $n \geqq 2$ such that $x^n = x$ identically, then the characteristic of R is square-free and is a divisor of the g.c.d. of $\{k^n - k : k = 1, 2, \ldots\}$. In *Certain congruences that hold identically* (this MONTHLY, 83(1976) 270–271) E. Hewitt proves the following (with $l+1$ replacing n): Let $m > 1$. Then there

exists an integer $n \geqq 2$ such that (i) $k^n \equiv k \bmod m$ for all integers k, if and only if m is square-free. For square-free m, write $m = p_1 \cdots p_r$, a product of distinct primes. Then the smallest $n \geqq 2$ such that (i) holds identically is $1 + \text{l.c.m.}\ \{p_1 - 1, \ldots, p_r - 1\}$. In *Multiplicatively periodic rings* (this MONTHLY, 83 (1976) 547–549) T. Chinburg and M. Henricksen characterize pairs of positive integers (n, m), $n \geqq 2$, such that $x^n \equiv x$, $mx = 0$ (m minimal for these properties) can hold identically in a ring. Further references can be found in the aforementioned articles.

General. Concerning his article *Variations on van der Waerden's and Ramsey's theorems* (this MONTHLY, 82(1975) 993–995) T. C. Brown writes that T. Kano has informed him that his Theorem V′ was proved previously by S. Kakeya and S. Morimoto (*On a theorem of M. Baudet and van der Waerden*, Jap. J. Math., 7(1930) 163–165). The latter article contains the following consequence of Theorem V′: There exists a function f such that $f(n) \to \infty$ as $n \to \infty$ and such that if $a_1 < a_2 < \cdots$ is any sequence of positive integers with $a_{n+1} - a_n \leqq f(n)$, then $\{a_i\}$ contains arithmetic progressions of arbitrary finite lengths.

R. D. Nelson has written us to say that correction (2) given in the last Comments and Complements article (83(1976) 798–801) to B. Fisher's *On a problem of Besicovitch* (this MONTHLY, 80(1973) 785–787) is itself in need of a correction. It should read

(2) replace line 13 by: $\frac{S}{2^{n-1}}\left(1 - x + \frac{x^2}{4}\right) < \frac{S}{2^{n-1}}(1 - \frac{1}{2}x)$.

In *An ancient unfair game* (this MONTHLY, 83(1976) 623–625). R. Feinerman shows that the game of dreydel is unfair. A simple derivation of the main theorem of that article, as well as a modification in the payoffs to make it fair, is given in an unpublished note of U. Yechiali entitled *How to make an ancient unfair game fair.* Yechiali proposes to let a player add M units (in place of 1) to the pot whenever the outcome is S. Then let Y_n be the amount in the pot just prior to the nth spin. Then Y_{n+1} will be p, $Y_n/2$, Y_n, or $Y_n + M$ accordingly as the outcome is G, H, N, or S, each with probability 1/4 (recall p is the number of players). Hence $E(Y_{n+1}) = \frac{1}{4}(p + M) + \frac{5}{8}E(Y_n)$, and since $E(Y_1) = p$,

$$E(Y_{n+1}) = \tfrac{1}{4}(p + M)\sum_{i=0}^{n-1}\left(\frac{5}{8}\right)^i + \left(\frac{5}{8}\right)^n p = \tfrac{2}{3}(p + M) + \left(\frac{p - 2M}{3}\right)\left(\frac{5}{8}\right)^n.$$

The payoff X_n on the nth spin equals Y_n, $Y_n/2$, 0, or $-M$ with equal probabilities, and hence

$$E(X_n) = \tfrac{1}{4}(\tfrac{3}{2}E(Y_n) - M) = \frac{p}{4} + \left(\frac{p - 2M}{8}\right)\left(\frac{5}{8}\right)^{n-1}.$$

For $M = 1$, Feinerman's theorem results; for $M = p/2$, the game is fair.

Concerning J. M. Weinstein's *Crowded posets and Ramsey's theorem* (this MONTHLY, 83(1976) 625–627) E. S. Wolk has written us that the sufficiency part of Theorem 4 is equivalent to the known result that the product of partially-well-ordered sets is partially-well-ordered. This theorem is proved by E. Michael in his article *A class of partially ordered sets* (this MONTHLY, 67(1960) 448–449) and was stated and used in Wolk's article *Partially well ordered sets and partial ordinals* (Fundamenta Mathematicae, 60(1967) 175–186). Michael establishes four equivalent definitions for a partially-well-ordered set, one of which is that there be no infinite descending chain and no infinite antichain. It should be noted that Weinstein defines a partially-well-ordered set by the property that it have no infinite descending chain. Thus a crowded partially-well-ordered set in the sense used by Weinstein is a partially-well-ordered set in the sense used by Michael.

L. Kuipers has written to tell us that the proof of L. Moser that the series of the reciprocals of the primes diverges (this MONTHLY, 65 (1958) 104–105), with very little modification gives a proof that the series of the reciprocals of the primes, congruent to 3 modulo 4, diverges. After replacing the sequence (p) of primes by the sequence (p) of primes congruent to 3 modulo 4, the only observation needed is that (referring to Moser's proof) since $T_i \equiv 3 \pmod 4$, T_i is divisible by at least one prime factor $p^* \equiv 3 \pmod 4$. For this p^*, $n!\,m > p^* > n$.

R. J. Evans has located an error in H. Harborth's article *Divisibility of binomial coefficients by their row number* (this MONTHLY, 84(1977) 35–37) which Harborth has corrected as follows: In equation (5) add: and $\binom{m}{k} \equiv 0 \pmod{t}$. The rightmost denominator in (7) should be $it + j$. In the last 7 lines of page 36 use t only for prime powers p^α, and write "there are *at most* $Z([n/t])\ldots$". In equation (9) write $2Z(\cdots)$ instead of $Z(\cdots)$. (Note that in (9) it suffices to subtract $2Z(\cdots)$ for all $t = p^x$.)

RESEARCH PROBLEMS

EDITED BY RICHARD GUY

In this Department the Monthly presents easily stated research problems dealing with notions ordinarily encountered in undergraduate mathematics. Each problem should be accompanied by relevant references (if any are known to the author) and by a brief description of known partial results. Manuscripts should be sent to Richard Guy, Department of Mathematics and Statistics, The University of Calgary, Calgary, Alberta, Canada, T2N 1N4.

MONTHLY RESEARCH PROBLEMS, 1969–77

RICHARD K. GUY

Our predictions that the Four Color Conjecture would be settled [1971, 1113] and that the solution would be unsuited to these pages [1973, 1121] have been fulfilled. For reasons discussed in [1971, 1114] and [1973, 1120] we don't publish *solutions* (exceptionally a few appear as Notes) but updating articles in December issues of odd-numbered years give information on progress with published Research Problems. Correspondence concerning policy, and problems published earlier, is welcomed. A recent suggestion is that we should accommodate less original papers commenting on classical open problems: the Riemann hypothesis, $P = NP$, the Poincaré conjecture.

References in brackets are years of updating articles [1971, 1973, 1975] and pages of this MONTHLY. References in parentheses are to articles listed at the end: years, tbp (to be published) or wrc (written communication).

In connexion with Klee's problem [1969, 286] on the expected volume of a simplex with vertices chosen randomly from a convex body, Groemer (1973, 1974) has shown this is minimal just if the body is an ellipsoid; he extends this result to higher order moments.

Thomassen (1976) throws further light on Kronk's problem [1969, 809; 1973, 1121; 1975, 996]: he finds infinite classes of planar hypohamiltonian graphs of connectivity 2 and girths 3, 4 and 5, and planar hypotraceable graphs of connectivity 3 and same girths. He asks: Does every finite hypohamiltonian graph contain a trivalent vertex? Is there a finite planar cubic hypohamiltonian (hypotraceable) graph? Is there an infinite, locally finite, hypohamiltonian (hypotraceable) graph? We now refer exactly to the paper of Doyen and Van Diest (1975). Doyen writes that his problem 3.7 has been solved by his student Simone Gult: the graphs $G_t\,(m, n)$ are hypohamiltonian for every odd $t \geqq 7$ provided m, n are not both 1. He reports a computer proof by J. B. Collier and E. F. Schmeichel (U.S.C.) of the non-existence of a hypohamiltonian graph of order 14 or a cubic hypohamiltonian graph of order 16. Related papers include Reid (1972) and Thomassen (1974).

Papers on Kronk's second problem [1969, 1045; 1971, 1116; 1973, 1121; 1975, 996] earlier given no reference, or inexact ones, are Bhat (1972), Bondy and Ingleton (1976), Chvátal (1973), Faudree and Schelp (1975–76), Fleischner (1975), Fleischner and Hobbs (1975), Häggkvist and Thomassen (1976) and Hakimi and Schmeikel (1974). Hobbs (1976) shows that the square of a graph is vertex-pancyclic; Faudree and Schelp (1976) improved this, showing that the square of a block is strongly path-

connected. Fleischner (1976) has further shown that in the square of a graph the following pairs of properties are equivalent: hamiltonian and hamiltonian connected; panconnected and pancyclic.

Ringel's problem, considered by Duke [1969, 1128] concerns a **proper labelling (graceful numbering**, Golomb; **β-valuation**, Rosa) of a graph on m vertices with $m-1$ edges, i.e., a labelling of the vertices with $[1, m]$ and edges with $[1, m-1]$ so that each edge-label is the difference of its vertex-labels. Rosa called a graph and its labelling **balanced** if there is an integer r so that the vertex-labels x, y of each edge satisfy $x \leqq r < y$. Sheppard (1976) calls $\{j_i\}$ a **labelling sequence** if $1 \leqq j_i \leqq m-i$ for $1 \leqq i \leqq m-1$; gives a $1-1$ correspondence between labelling sequences and graphs with a proper labelling; and enumerates balanced labelled graphs. Kotzig (1975) has studied valuation of graphs whose components are circuits. See also [1973, 1122; 1975, 996], Kotzig (1973), Gyárfás and Lehel (tbp) and comments on [1974, 499] below.

Wall (1975) establishes that his unitary perfect number [1970, 389] is the fifth such.

We give the exact reference to Carlitz's proof (1971) of Gandhi's conjecture [1970, 505; 1971, 1117; 1975, 996] on Genocchi numbers.

Lawrence (1977) has settled a conjecture of Grünbaum [1970, 1088]; he shows that if $d \geqq 7$, $e \geqq 4$, there is no d-valent graph of chromatic number d and girth e.

McMullen's solution (1974) of Rosenfeld's problem [1971, 49; 1971, 1119; 1973, 1124; 1975, 997] for the case of 3 cliques has appeared; and see his paper with Altschuler (1973). Rosenfeld mentions connexions between his problem and strong products of graphs (Rosenfeld, 1967, 1970), the perfect graph conjecture (Chvátal, 1975) and integer programming (Padberg, 1974).

An exact reference is at last given to Singmaster's paper (1975) on his own problem [1971, 385; 1971, 1119; 1973, 1124; 1975, 997].

Schneider (1975) disproves a conjecture of Bolker [1971, 529; 1973, 1124; 1975, 997] by showing that Euclidean n-space contains non-ellipsoidal, centrally symmetric convex bodies which, as well as their polars, are zonoids. Dor (1976) settles another question of Bolker by showing (via the dual result that l_n^p does not embed isometrically in L_1 for $n \geqq 3$, $p > 2$) that the unit ball of l_n^q is not a zonoid for $n \geqq 3$, $1 \leqq q < 2$.

Dekking (1976) and Doyle (tbp) use different methods to refine the discoveries of Thue, mentioned by Brown [1971, 886] and of Entringer, Jackson and Schatz [1975, 997]: there are infinite binary sequences with no 3 adjacent identical blocks and no 2 adjacent identical blocks of length 4 or more; binary sequences without 3 adjacent identical blocks and no adjacent blocks of length 3 or more are finite; infinite binary sequences that have no identical overlapping blocks have arbitrarily long identical adjacent blocks. Fife (1976) shows the essential difference between 1-sided and 2-sided sequences that contain no overlapping blocks $b_1b_2\ldots b_nb_1b_2\ldots b_nb_1$. Stop press: Dekking has solved the (3,3) and (2,4) problems, and begins to doubt Brown's "duality" principle.

The correct references to Herda's extensions (1975, 1976) to 3 dimensions of his 2-dimensional problem [1971, 888; 1971, 1120; 1973, 1124; 1975, 997] are now available, as is that to Rowen's paper (1974) settling the conjecture of Smith and Kumin [1972, 157; 1973, 1124; 1975, 997].

Goodey and Woodcock (tbp) have solved Peterson's problem [1972, 505] on sets of constant width.

Mather (1976) has Biggs's footballers [1972, 1018; 1973, 1125; 1975, 998] playing rugger: he finds with computer help a hamilton circuit of O_8, whose $\binom{15}{7}$ vertices are the 7-subsets of a 15-set, joined just if the subsets are disjoint. Can O_8 be partitioned into 4 edge-disjoint hamilton circuits? See also de Werra (1975).

Pomerance's papers (1975, 1975–76, 1977) on Lehmer's problem (Alter [1973, 192; 1975, 998]) have appeared. Arthur Marshall draws attention to the solution by O. P. Lossers to Problem 721 in Élem. Math., 30 (1975) 88 and the use of Pólya's theorem (1925).

J. van Leeuwen reports progress on Nash's reachability problem [1973, 292]: he solved [1975, 1003] the general problem in dimension $\leqq 3$. Hopcroft and Pansiot (1976) extended his approach to $d \leqq 5$; Sacerdote and Tenney (1977) found a complete decision procedure in any dimension. Their proof is "complicated, technical, not quite fully understood by most of us" but there's general agreement the

problem is settled affirmatively. Equivalent problems are treated by Hack (1974, 1975) and Araki and Kasami (tbp); related papers by Anshel (1976), H. Baker (1973), Cardosa (1975) and Lipton (1975).

Joel Spencer writes that Fristedt (tbp) has solved his deception game [1973, 416]. The answer is yes; the changer can hold the chooser to expectation $\frac{1}{2}$, which is what he gets by choosing randomly: this holds for any symmetric initial distribution.

If Ω_n is the multiplicative semigroup of doubly stochastic matrices of order n, Butler (1975) calls $A \in \Omega_n$ **regular** if $\exists B \in \Omega_n \ni A = BAB$, and proves van der Waerden's conjecture [1973, 791; 1975, 998], $\operatorname{per} A \geqq n!/n^n$ for regular A. Sinkhorn (wrc) proves the same result by purely matrix methods. Gyires (1976) proves $(\operatorname{per} AA^T + \operatorname{per} A^TA + 2\operatorname{per} A^2)/4 \geqq n!/n^n$ and conjectures $4(\operatorname{per} A^2)/(\operatorname{per} AA^T + \operatorname{per} A^TA + 2\operatorname{per} A^2) \geqq n!/n^n$. See also Henderson (1975).

David Zeitlin (wrc) makes extensive comments and asks numerous questions concerning the U-numbers discussed by Recamán [1973, 919; 1975, 998]. He defines $a(n)$, $b(n)$ by $a(0) = 0 = b(0)$, $u_{n+3} = u_{n+2-a(n)} + u_{n+1-b(n)}$, $n \geqq 0$.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23...
u_n	1	2	3	4	6	8	11	13	16	18	26	28	36	38	47	48	53	57	62	69	72	77	82...
$a(n)$	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	5	2	0	1	2	3	4	0...
$b(n)$	1	1	2	2	4	4	6	3	8	5	10	6	13	10	12	6	9	16	14	13	12	11	22...

He notes that $b(n) = \phi(n)$, Euler's totient function, for $n \leqq 5$, $n = 7$, 11, 21, 23 and asks if this happens infinitely often. He proves (1975) that $\{u_n\}$ is **complete** in that every positive integer is the sum of distinct U-numbers. He asks if the Fibonacci and Lucas numbers are always the sum of at most 2 U-numbers. He notes that $u_n(u_{n-1} + u_{n+1})$ is a U-number for $2 \leqq n \leqq 5$ and asks if this occurs infinitely often. Such events may best be attributed to the strong law of small numbers.

Bollobás and Guy (tbp) have settled Meyer's problem [1973, 920; 1975, 998–999]: trees on n vertices with maximum valence Δ are equitably 3-colorable if $n = \max(\Delta + 2, 3\Delta - 10)$ or $n \geqq \max(\Delta + 1, 3\Delta - 8)$. More generally, equitably k-colorable, $k \geqq 3$, provided that if $\Delta = (q+2)(k-2) + r$ with $3 \leqq r \leqq k$, $n \geqq \max(\Delta + 1, \Delta + 2q + 2)$ or $r = 3$, $q > 0$, $n = \Delta + 2q$. There are corresponding results for **proportional coloring**, provided not more than half the vertices are required to be in one class.

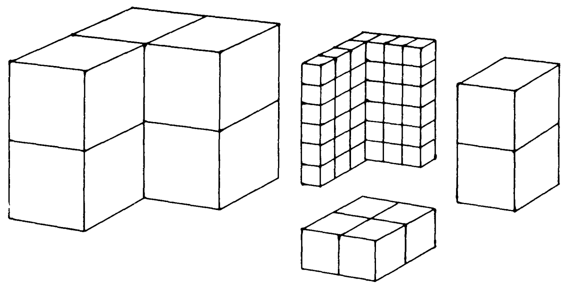
Deakin writes about his problem [1974, 56; 1975, 999] that a numerical counterexample, both to his own and to Feldman's versions of the conjecture, was found by Karlin and Carmelli (1975). Characterization of "fitness matrices" which lead to this counterintuitive situation remains an open question.

Molnar forwards further solutions to his matrix problem [1974, 383; 1975, 999–1000] sent by Francis Coghlan, The University, Manchester, M13 3PL, England and Kenneth Lau, Central College, Pella, IN 50219:

$$\begin{bmatrix} 3 & 9 & 11 \\ 2 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix} \qquad \begin{bmatrix} 3 & 2 & 3 \\ 5 & 3 & 2 \\ 7 & 5 & 9 \end{bmatrix}$$

Golomb showed that K_n can be gracefully numbered just if $n < 5$ and Kotzig and Turgeon (1975) that cK_n can be just if $c = 1$ and $n < 5$. Other papers on Golomb's problem [1974, 499] of the largest graceful subgraph, and related problems of difference bases and economical rulers, not mentioned in [1975, 1000] are Bloom and Golomb (1976, tbp), Brauer (1945), Haselgrove and Leech (1957) and Piccard (1939). The IEEE paper of Bloom and Golomb has an extensive bibliography and see the remarks on Ringel's problem [1969, 1128] above and Cahit's [1976, 35] below.

Chr. Meier writes about his submission [1974, 630; 1975, 1001] on Hadwiger's cube decomposition problem: the dissection into 54 cubes was by Miss Doris Rychener, teacher of the flute at the *Konservatorium für Musik*, Bern, and independently by A. Zbinden of IBM, Bern:



Erdös (1974) uses a theorem of A. Brauer to improve the bound on $c(n)$ from 2^{n^2} to $(e-1)(2n)^n$.

The second paper of Chinburg and Henriksen (1976), relevant to Edgar's exponential diophantine equation problem [1974, 758; 1975, 1000], has appeared, as has that of Tung-Po Lin (1977) settling P. R. Scott's area-perimeter problem [1974, 884; 1975, 1001].

Loveland writes concerning his paper with Wayment [1974, 1003; 1975, 1001]: the conjecture and four questions are still open, but the conjecture is true if the midset function is continuous (1976a). He has generalized the midset idea to Apollonius- or λ-sets and shown that a continuum with the double λ-set property and continuous λ-set function is a simple closed curve (1976b). With Valentine (1975) he gave a short proof that a complete convex non-degenerate metric space with the double midset property is isometric to a convex circle. See also Berard and Nitka (1974).

Witsenhausen writes that the conjecture in his paper [1974, 1100] was stated for $d=2$ by Kelly (1969) who proves it in this case for $n \leqq 9$ by *ad hoc* methods. Kelly's conjecture that for $d=2$, $n \geqq 6$, the sum of distances is maximized for given diameter by distributing the points as evenly as possible among the vertices of an equilateral triangle, is false; Witsenhausen can show it can hold only for finitely many n. Larman and Rogers (1972) had asymptotically improved the upper bound in his second problem [1974, 1101] to $(\log_2 n)^3/n^2$.

The papers of Maria Klawe (tbp) and Kermit Sigmon (tbp) contain counterexamples to Chow's conjecture [1975, 155; 1975, 1001].

In answer to question 2 in §7 of D'Alarcao and Moore [1975, 270] S. Baskaran (tbp) proves that a finite soluble group has all its Sylow subgroups cyclic just if the orders of the intersection and lattice-union of any 2 subnormal subgroups are the g.c.d. and l.c.m. of the orders of the individual groups.

Pomerance notes that Eggleton [1975, 499; 1975, 1001] did not mention **locally finite tilings** (compact sets intersect finitely many tiles). His solution (1977) gives such a tiling, but uses the Nagell–Lutz theorem on elliptic curves. He obtains a more elementary solution by sacrificing local finiteness. He and Eggleton (tbp) prove that there is a (not locally finite) strict 3-space tiling with rational edged tetrahedra, one from each isometry class.

Problem 1 of Alpert and Gross [1975, 835] attributed to R. H. Fox and given as unsolved in Ringel's book (1975, p. 179) had been settled earlier, correspondents observed. Grünbaum and Motzkin (1963) solved the problem in dual formulation. See §13.4, Theorems 1 and 2, of Grünbaum's book (1967). Feser notes that a proof of Theorem 1 is in his dissertation (1975). Zaks gave the Grünbaum and Ringel references and noted that Étourneau (1973) proved Theorem 2. Malkevitch adds further references: Grünbaum (1968), Grünbaum and Zaks (1974), Jucovič (1971) and Malkevitch (1969, 1970). Mitchem (1976) gave counterexamples.

Andrews sends a paper by Porubský (tbp). His problem [1975, 922] suggested several problems to Erdös: If $0<x_1<x_2<\cdots$ is an arbitrary integer sequence with no x_n of form $x_u+x_{u+1}+\cdots+x_v$, is the lower density 0? The upper density can be 1/2 but probably not more. The logarithmic density is probably 0, i.e., $(\Sigma\, 1/x_i)/\ln n \to 0$, where the sum is for $x_i<n$. Let $1 \leqq a_1 < \cdots < a_r \leqq n$. Denote by $A(n)$ the least integer such that for every choice of a's, $\exists m$ with $n<m\leqq A(n)$, $m \neq a_u+a_{u+1}+\cdots+a_v$; is it true that $A(n)<Cn$ for some C?

Assume now that all sums $a_u + a_{u+1} + \cdots + a_v$ are distinct; is $\Sigma_{a_i<n} 1 = o(n)$? Is $\Sigma 1/a_i < \infty$? Long ago Leo Moser and Erdös considered sums of consecutive primes $p_u + p_{u+1} + \cdots + p_v$ but could prove nothing. Is the density of such sums positive? Is the number of solutions of $n = p_u + p_{u+1} + \cdots + p_v$ bounded? Is there an integer sequence $a_1 < a_2 < \cdots$ so that all but a finite number of n are of the form $n = a_u + a_{u+1} + \cdots + a_v$, $v > u$? (If $a_i = i$, 2^k is not of this form.)

Write $\Pi(u, v) = u(u+1)\cdots v$ and let $F(n)$ be the number of solutions of $\Pi(u_1, v_1) = \Pi(u_2, v_2)$ with $v_1, v_2 < n$. Since $u(u+1) = u^2 + u$, $F(n) > \sqrt{n}$. How small is $F(n) - \sqrt{n}$? Let $a_1 < a_2 < \cdots$ be a sequence of *reals*, $A(u, v) = a_u + a_{u+1} + \cdots + a_v$ and suppose $|A(u_1, v_1) - A(u_2, v_2)| > c$; prove that $\Sigma 1/a_n < \infty$ and that $a_n/n \to \infty$.

Herbert Taylor (1977) neatly answers a question about Leech's tree edge-labelling problem [1975, 923] by 2-coloring the n vertices consistently with even paths joining like colors, odd joining unlike: $b + w = n$, $2bw = \binom{n}{2}$ or $\binom{n}{2} + 1$, so $n = k^2$ or $k^2 + 2$.

All complete binary trees *are* graceful [1976, 35]: we apologize for not noting that this follows from results of Stanton and Zarnke (1973) quoted in [1973, 1122; 1975, 996]; hopefully Brualdi's note [1976, 801] reduced further duplication of effort. Joel Brenner observed misprints in the Figures: in Fig. 2 the entry 4 is missing from position 1, 0 in the 15-tree; in the 31-tree the entry 2 at 0, 0, 1 should be 9 and 17 at 1, 1, 0, 1 should be 22: and noted the ambiguity of "these" on line 1 of p. 37; the intention was to include all trees obtained from those given, by either subtracting all labels from 2^k or interchanging two subtrees dependent on a non-terminal vertex. Barnes (wrc) proved that label $2^k - 1$ could always appear at the apex (or root, of trees not growing in Australia). Gary Bloom reported that Herb Taylor had solved the problem and that Osterweil had given him a copy of Gabow's paper (1975) which acknowledges Osterweil's help. Max Warschauer sent a solution and Owens (1976) showed more generally that complete d-nary trees are graceful. This is also one of the results of Gyárfás and Lehel (tbp).

C. E. H. Francis (wrc), Jesus College, Cambridge, cuts a strip from the top of the "Shephard piano" and extends the base to obtain a sofa of area 2.215281, larger than that of Wagner [1976, 188]. Iteration further improves this to 2.21563. The present writer also found an area 2.21503; by combining ideas an area 2.215649 is achieved, likely to be near that of the best sofa, whose existence, if not its symmetry or smoothness, is proved by Mike Guy (wrc) and John Conway (wrc).

Alan Baker observes that Wilansky's question [1976, 473] is answered if Schanuel's conjecture is true—see his book (1975)—but this is a hard nut to crack. Wagner (wrc) used a computer to show that if 2^c, 3^c are integers, but c is not, then $2^c > 10^7$.

Groemer (tbp) sends preprints, one with Chakerian (tbp), on his covering and packing problems [1976, 726].

Ronald Evans notes that Rosen's final conjecture [1977, 39] is false and quotes Borho and Rosenberger (1973), while Roger Lyndon notes relations between the problem of Brenner and Riddell [1977, 39] and questions in Magnus (1974) and Tretkoff (1975).

Hazelgrove and Leech write that Robert E. Kibler observes that in their tournament design problem [1977, 198] for $2n = 8$ players there are 8 court allocations (4 symmetrical) not 6 (2 sym.) as stated. Paul Smith (tbp) has a solution for some outstanding values of $2n$, including 22, 34, 46, 58 (mod 60); he relates the problem to double Howell rotations and skew Room squares, constructible from results of Mullin and Nemeth (1969), Wallis (1972, 1974, et al. 1972), Dillon and Morris (1973) and Stanton and Horton (1972). The solutions satisfy the extra constraint of Gelling and Odeh (1973) that no player plays consecutively on the same court. Mullin and Wallis (1975) give a Room square bibliography. For a related problem see de Werra (1975). Stop press: Schellenberg *et al* (1977) have solved the problem.

In the theorem of Mycielski [1977, 117] proposition (b) should read "For all non-empty sets P, $Q \in \boldsymbol{B}^*$ we have $P \cong Q$."

Michael Patterson (wrc), Floyd (tbp) and Fredman (tbp) have settled Klee's question [1977, 284].

As usual I thank numerous correspondents, not all mentioned here, who have sent comments,

references, offprints and preprints. The value of these articles is largely due to your efforts; keep up the good work!

References

A. Altschuler and P. McMullen, The number of simplicial neighbourly d-polytopes with $d+3$ vertices, Mathematika, 20 (1973) 263–266; MR **50** # 1124.

Michael Anshel, Decision problems for HNN groups and vector addition systems, Math. Comput., 30 (1976) 154–156; MR **53** # 626.

T. Araki and T. Kasami (tbp) Some decision problems related to the reachability problem for Petri-nets, Theor. Comp. Sci.

A. Baker, Transcendental Number Theory, Cambridge Univ. Press, 1975, pp. 119–120.

Henry G. Baker, Rabin's proof of the undecidability of the reachability set inclusion problem for vector addition systems, Computation Structures Group Memo 79, Project MAC, M.I.T., July 1973.

Wilfred E. Barnes (wrc) Every complete binary tree is graceful.

S. Baskaran (tbp) A characterization of finite groups with cyclic Sylow subgroups, Math. Student,

A. D. Berard and W. Nitka, A new definition of the circle by the use of bisectors, Fund. Math., 85 (1974) 49–55; MR **50** # 8464.

Vasanti N. Bhat, The powers of a connected graph are highly hamiltonian-connected, Indian J. Pure Appl. Math., 3 (1972) 1247–1249; MR **50** # 9657.

——— and S. F. Kapoor, Minimally n-Hamiltonian graphs, J. Indian Math. Soc. (N.S.), 37 (1973) 277–284 (1974); MR **53** # 2743.

Gary S. Bloom, A counterexample to a theorem of S. Piccard, J. Combinatorial Theory Ser. A, 22 (1977) 378–379.

Gary S. Bloom and Solomon W. Golomb, Numbered complete graphs, unusual rulers and assorted applications, Theory and Applications of Graphs in America's Bicentennial Year (Proc. Conf. Graphs, Kalamazoo, Michigan, May 1976) Springer Lecture Notes (1977).

——— and ——— (tbp) On the largest graceful subgraph of the complete graph,

——— and ——— (tbp) Numbered undirected graphs and their applications—a unifying model, Proc. I.E.E.E.

B. Bollobás and Richard K. Guy (tbp) Equitable and proportional coloring of trees, Discrete Math.,

J. A. Bondy and A. W. Ingleton, Pancyclic graphs II, J. Combinatorial Theory Ser. B, 20 (1976) 41–46; MR **53** # 5366.

W. Borho and G. Rosenberger, Eine Bemerkung zur Hecke-Gruppe $G(\lambda)$, Abh. Math. Sem. Univ. Hamburg, 39 (1973) 83–87; MR **48** # 8391.

A. Brauer, A problem of additive number theory and its application to electronic engineering, J. Elisha Mitchell Sci. Soc., 61 (1945) 55–66.

Kim Ki-Hang Butler, On the van der Waerden permanent conjecture, Abstract 75T–A234, Notices Amer. Math. Soc., 22 (1975) A–620.

E. W. Cardoza, Computational complexity of the word problem for commutative semi-groups, MAC Tech. Memo. 67, M.I.T., Oct. 1975.

L. Carlitz, A conjecture concerning Genocchi numbers, Norske Vid. Selsk. Skr. (Trondheim) 9 (1971) 1–4; MR **45** # 6749.

G. D. Chakerian and H. Groemer (tbp) On coverings of euclidean space by convex sets, Pacific J. Math.,

Ted Chinberg and Melvin Henriksen, Sums of kth powers in the ring of polynomials with integer coefficients, Acta Arith., 29 (1976) 227–250; MR **53** # 7942.

Václav Chvátal, New directions in hamiltonian graph theory, New Directions in the Theory of Graphs (Proc. 3rd Ann Arbor Conf. Graph Theory, 1971) Academic Press, 1973, 65–95; MR **50** # 9689.

———, On certain polytopes associated with graphs, J. Combinatorial Theory Ser. B, 18 (1975) 138–154; MR **51** # 7949.

F. M. Dekking, On repetitions of blocks in binary sequences, J. Combinatorial Theory Ser. A, 20 (1976) 292–299.

J. F. Dillon and R. A. Morris, A skew Room square of side 257, Utilitas Math., 4 (1973) 187–192; MR **48** # 3766.

Leonard E. Dor, Potentials and isometric embeddings in L_1, Israel J. Math., 24 (1976) 260–268.

———, On isometric embeddings in L_1, Article VII, The Altgeld Book, Univ. Illinois Functional Anal. Sem., Urbana, 1976.

Jean Doyen and Viviane van Diest, New families of hypohamiltonian graphs, Discrete Math., 13 (1975) 225–236.

——— and ———, Graphes hypohamiltoniens, Cahiers Centre Études Recherches Opér., 17 (1975) # 2–3–4, 191–192; MR **52** # 13496.

Jon Doyle (tbp) Non-repetitive binary sequences,

P. Erdös, Remarks on some problems in number theory, Conf. of Balkan Math., Belgrade, June 1974.

E. Étourneau, Existence and connectivity of planar graphs having 12 vertices of degree 5 and $n-12$ vertices of degree 6, A. Hajnal et al (eds.) Infinite and Finite Sets, Colloq. Math. Soc. J. Bolyai **10**, Keszthely, Hungary, 1973, North-Holland, 1975, 645–655; MR**52** # 158.

R. J. Faudree and R. H. Schelp, The entire graph of a bridgless connected plane graph is panconnected, J. London Math. Soc. (2), 12 (1975–76) 59–66; MR**52** # 5460.

——— and ———, The square of a block is strongly path connected, J. Combinatorial Theory Ser. B, 20 (1976) 47–61.

Earl Dennet Fife, Binary sequences which contain no BBb, PhD thesis, Wesleyan Univ., Middletown, Connecticut, Sept. 1976.

Herbert Fleischner, In the square of a graph, hamiltonianity and pancyclicity, hamiltonian connectedness and panconnectedness are equivalent concepts, Monatsh. Math., 82 (1976) 125–149.

———, Hamiltonsche totale Graphen von ebenen Graphen, Math. Nachr., 68 (1975) 83–91; MR**52** # 5476.

——— and A. M. Hobbs, A necessary condition for the square of a graph to be hamiltonian, J. Combinatorial Theory Ser. B, 19 (1975) 97–118; MR**54** # 2535.

——— and ———, Hamiltonian total graphs, Math. Nachr., 68 (1975) 59–82; MR**52** # 5475.

Robert W. Floyd (tbp) The measure of $U_1^n[a_i, b_i]$ cannot be computed in less than $O(n \log n)$ steps.

Michael L. Fredman (tbp). A note on the complexity of determining the measure of $\cup_{i=1}^{n}[a_i, b_i]$, Comm. ACM,

Bert Fristedt (tbp) Deception game of Thompson,

Harold N. Gabow, How to gracefully number certain symmetric trees, SIGACT News, 7 # 4 (Nov.-Dec. 1975) 33–36.

E. N. Gelling and R. E. Odeh, On 1-factorizations of the complete graph and the relationship to round robin schedules, Proc. 3rd Manitoba Conf. Numerical Math., Winnipeg (1973) 213–221; MR**50** # 171.

P. R. Goodey and M. M. Woodcock (tbp) A characterization of sets of constant width.

H. Groemer, On some mean values associated with a randomly selected simplex in a convex set, Pacific J. Math., 45 (1973) 525–533; MR**47** # 5916.

———, On the mean value of the volume of a random polytope in a convex set, Arch. Math. (Basel) 25 (1974) 86–90; MR**49** # 6036.

——— (tbp) On finite classes of convex sets that permit space coverings,

Branko Grünbaum, Convex Polytopes, Interscience, New York, 1967.

———, Some analogues of Eberhard's theorem on convex polytopes, Israel J. Math., 6 (1968) 398–411; MR**39** # 6168.

——— and T. S. Motzkin, The number of hexagons and the simplicity of geodesics on certain polyhedra, Canad. J. Math., 15 (1963) 744–751; MR**27** # 4133.

——— and Joseph Zaks, The existence of certain planar maps, Discrete Math., 10 (1974) 93–115; MR**50** # 1949.

A. Gyárfas and J. Lehel (tbp) A method to generate graceful trees, Proc. Colloq. Internat. C.N.R.S., Paris, July 1976.

B. Gyires, On a conjecture of van der Waerden type, Abstracts of Talks, 5th. Hungar. Combinatorial Colloq., Keszthely, 1976, 73.

Michel Hack, The recursive equivalence of the reachability-problem and the liveness problem for vector addition systems and Petri-nets, Conf. Rec. 15th IEEE Symp. Switching and Automata Theory (1974) 156–164.

———, Decision problems for Petri nets and vector addition systems, Computation Structures Group Memo 95, Project MAC, M.I.T., March 1975.

Roland Häggkvist and Carsten Thomassen, On pancyclic digraphs, J. Combinatorial Theory Ser. B, 20 (1976) 20–40; MR**52** # 10481.

S. L. Hakimi and E. F. Schmeichel, Pancyclic graphs and a conjecture of Bondy and Chvátal, J. Combinatorial Theory Ser. B, 17 (1974) 22–34.

C. B. Haselgrove and J. Leech, Note on restricted difference bases, J. London Math. Soc., 32 (1957) 228–231; MR**19**, 942.

J. R. Henderson, Permanents of (0, 1) matrices having at most two zeros per line, Canad. Math. Bull., 18 (1975) 353–358; MR**53** # 2978.

Hans Herda, A characterization of spheres among convex 3-bodies, Colloq. Math., 33 (1975) 103–104; MR**53** # 9040.

———, A characterization of spheres among compact 3-bodies satisfying Crofton's theorem, Colloq. Math., 36 (1976) 235–236.

Arthur M. Hobbs, The square of a block is vertex pancyclic, J. Combinatorial Theory Ser. B., 20 (1976) 1–4.

John Hopcroft and Jean-Jacques Pansiot, On the reachability problem for 5-dimensional vector-addition systems, TR76-280, Dept. Comp. Sci., Cornell Univ., 1976.

Ernest Jucovič, On the number of hexagons in a map, J. Combinatorial Theory Ser. B, 10 (1971) 232–236; MR**43** # 3910.

Samuel Karlin and Dorit Carmelli, Numerical Studies on two-loci selection models with general viabilities, Theor. Pop. Biol., 7 (1975) 399–421; MR**52** # 5081.

Leroy M. Kelly, Optimal distance configurations, Recent Progress in Combinatorics (Proc. 3rd Waterloo Conf. Combinatorics, 1968) Academic Press, 1969, 111–122; MR**40** # 7138.

Maria Klawe (tbp) A topologically simple semigroup that is not simple,

Anton Kotzig, On certain vertex-valuations of finite graphs, Utilitas Math., 4 (1973) 261–290; MR**52** # 5490.

———, β-valuations of quadratic graphs with isomorphic components, Utilitas Math., 7 (1975) 263–279; MR**52** # 2964.

——— and J. Turgeon, β-valuations of regular graphs with complete components, CRM-498, Montréal, Mar. 1975.

D. G. Larman and C. A. Rogers, The realization of distances within sets in Euclidean space, Mathematika, 19 (1972) 1–24; MR**47** # 7601.

Jim Lawrence, Covering the vertex set of a graph with subgraphs of smaller degree, Discrete Math., 21 (1978).

Tung-Po Lin, Maximal area under constraint, Math. Mag., 50 (1977) 32–34.

Richard J. Lipton, The reachability problem requires exponential space, Res. Report # 62, Dept. Comput. Sci., Yale Univ., January 1976.

L. D. Loveland, A metric characterization of a simple closed curve, General Topology and Appl., 6 (1976) 309–313; MR**53** # 9170.

———, Continuous finite Apollonius sets in metric spaces, Canad. J. Math., 28 (1976) 341–346; MR**53** # 6512.

——— and J. E. Valentine, Characterizing a circle with the double midset property, Proc. Amer. Math. Soc., 53 (1975) 443–444; MR**52** # 9079.

W. Magnus, Non-euclidean tesselations and their groups, Academic Press, 1974, §III.2, 112–122.

Joseph Malkevitch, Ph.D. thesis, Univ. of Wisconsin, 1969.

———, 3-valent 3-polytopes with faces having fewer than 7 edges, Annals N.Y. Acad. Sci., 175 (1970) 285–286; MR**42** # 114.

Michael Mather, The rugby footballers of Croam, J. Combinatorial Theory Ser. B, 20 (1976) 62–63; MR**52** # 10494.

P. McMullen, The number of neighbourly d-polytopes with $d+3$ vertices, Mathematika, 21 (1974) 26–31; MR**51** # 4054.

John Mitchem, Counterexamples to a conjecture by R. H. Fox, Abstract 76T–A94, Notices Amer. Math. Soc., 23 (1976) A-353–354.

R. C. Mullin and E. Nemeth, An existence theorem for Room squares, Canad. Math. Bull., 12 (1969) 493–497.

——— and W. D. Wallis, The existence of Room squares, Aequationes Math., 13 (1975) 1–7; MR**53** # 2715.

Jan Mycielski (tbp) Addendum to paper: Can one solve equations in groups?

Frank W. Owens, Complete d-nary trees are graceful, Abstract 76T–A97, Notices Amer. Math. Soc., 23 (1976) A-354.

Manfred W. Padberg, Perfect zero-one matrices, Math. Programming, 6 (1974) 180–196; MR**49** # 4809.

Sophie Piccard, Sur les ensembles de distances des ensembles de points d'un espace euclidien, Mémoires de l'Univ. de Neuchâtel, 13, Gauthier-Villars, Paris, 1939, esp. p. 31; MR**2**, 129–130.

G. Pólya and G. Szegö, Aufgaben und Lehrsätze aus der Analysis, Vol. 2, Springer, Berlin, 1925, pp. 133, 342.

Carl Pomerance, On the congruences $\sigma(n) \equiv a \pmod n$ and $n \equiv a \pmod{\varphi(n)}$, Acta Arith., 26 (1974–75) 265–272; MR**52** # 5535.

———, On composite n for which $\varphi(n) | n-1$, Acta Arith., 28 (1975–76) 387–389; MR**52** # 13608.

———, On composite n for which $\varphi(n) | n-1$, II, Pacific J. Math., 69 (1977) 177–186.

———, On a tiling problem of R. B. Eggleton, Discrete Math., 18 (1977) 63–70.

K. B. Reid, On hypotraceable graphs, Proc. 3rd S.E. Conf. Combinatorics, Graph Theory and Computing, Boca Raton, 1972, 409–422; MR**50** # 1972.

G. Ringel, Map Color Theorem, Springer-Verlag, New York, 1974.

M. Rosenfeld, On a problem of C. E. Shannon in graph theory, Proc. Amer. Math. Soc., 18 (1967) 315–319; MR**34** # 7405.

———, On certain polytopes associated with graphs, Proc. Louisiana Conf. Combinatorics, Graph Theory and Computing, March, 1970, 268–280; MR**42** # 1701 (& p. 1825).

Louis H. Rowen, Standard polynomials in matrix algebras, Trans. Amer. Math. Soc., 190 (1974) 253–284; MR**50** # 2208.

George S. Sacerdote and Richard L. Tenney, The decidability of the reachability problem for vector-addition systems, 9th ACM Symp. Theory Comput., (1977) 61–76.

P. J. Schellenberg, G. H. J. van Rees and S. A. Vanstone, The existence of balanced tournament designs, Ars Combinatoria, 3 (1977) 303–317.

Rolf Schneider, Zonoids whose polars are zonoids, Proc. Amer. Math. Soc., 50 (1975) 365–368.

David A. Sheppard, The factorial representation of balanced labelled graphs, Discrete Math., 15 (1976) 379–388.

Kermit Sigmon (tbp) On when a topologically simple semigroup is simple,

David Singmaster, Repeated binomial coefficients and Fibonacci numbers, Fibonacci Quart., 13 (1975) 295–298.

Richard Sinkhorn (wrc) Concerning regular doubly stochastic matrices.

Paul Smith (tbp) On double Howell rotations,

R. G. Stanton and J. D. Horton, Composition of Room squares, Combinatorial Theory and its Applications III (Proc. Colloq., Balatonfüred, 1969) North-Holland, 1970, 1013–1021; MR**46** # 73.

——— and ———, A multiplication theorem for Room squares, J. Combinatorial Theory Ser. A, 12 (1972) 322-325; MR**45** # 6650.

——— and C. R. Zarnke, Labelling of balanced trees, Congressus Numerantium VIII, Proc. 4th S.E. Conf. Combinatorics, Graph Theory and Computing, Boca Raton, 1973, 479–495.

Herbert Taylor, On the number of odd path sums in an edge labelled tree, Math. Mag., 50 (1977).

Carsten Thomassen, Graphs in which every path is contained in a Hamilton path, J. Reine Angew. Math., 268–269 (1974) 271–282; MR**50** # 1977.

———, Planar and infinite hypohamiltonian and hypotraceable graphs, Discrete Math., 14 (1976) 377–389.

Carol Tretkoff, Nonparabolic subgroups of the modular group, Glasgow Math. J., 16 (1975) 91–102; MR**53** # 3140.

Neal R. Wagner (wrc) A computer search for integer values of 2^c and 3^c.

Charles R. Wall, The fifth unitary perfect number, Canad. Math. Bull., 18 (1975) 115–122; MR**51** # 12690.

W. D. Wallis, Duplication of Room squares, J. Austral. Math. Soc., 14 (1972) 75–81; MR**47** # 3213.

———, Solution of the Room square existence problem, J. Combinatorial Theory Ser. A, 17 (1974) 379–383; MR**50** # 6899.

———, A. P. Street and J. S. Wallis, Combinatorics: Room squares, sum-free sets, Hadamard matrices, Springer-Verlag, 1972; MR**52** # 13397.

D. de Werra, On good and equitable edge-colorings, Cahiers Centre Études Recherche Opér., 17 (1975) # 2–3–4, 417–426.

———, On a particular conference scheduling problem, INFOR-Canad. J. Operational Res. and Information Processing, 13 (1975) 308–315; MR**52** # 16596.

David Zeitlin, Ulam's sequence $\{U_n\}$, $U_1 = 1$, $U_2 = 2$, is a complete sequence, Abstract 75T–A267, Notices Amer. Math. Soc., 22 (1975) A–707.

DEPARTMENT OF MATHEMATICS AND STATISTICS, THE UNIVERSITY OF CALGARY, CALGARY, ALBERTA, CANADA, T2N 1N4.

CLASSROOM NOTES

EDITED BY RICHARD A. BRUALDI

Material for this Department should be sent to Richard A. Brualdi, Department of Mathematics, University of Wisconsin, Madison, WI 53706.

BEYOND THE MYSTIC HEXAGRAM

MIGUEL DE GUZMÁN

The théorem of Brianchon states that if $V_1V_2V_3V_4V_5V_6$ is a hexagon whose sides are tangent to a conic C, then the three diagonals V_1V_4, V_2V_5, V_3V_6 meet at a point B.

An intuitive proof of this fact is presented in [**1**]. The following is a still simpler proof which leads in a natural way to the generalization of Brianchon's theorem stated at the end of this note.

It is clearly sufficient to assume that the conic C is a circle, since, by projection and section, any other situation can be reduced to this one. The proof of Brianchon's theorem is then a simple consequence of the two following observations.

OBSERVATION 1. Let π be the plane on which the circle C lies. Fix an orientation over its circumference Γ. Choose two arbitrary points P and Q on Γ. Starting at P draw two half lines, one the tangent p of Γ at P and having the same sense as Γ, and the other, the tangent q of Γ at Q and in the opposite sense as Γ.

Consider also a half line p_1 emanating from P and located on the plane that contains p and is perpendicular to π, and another half line q_1 emanating from Q and situated on the plane containing q perpendicular to π. Assume that p_1, q_1 are in the same half space determined by π and that the angles p_1Pp, q_1Qq are equal. Then the lines containing p_1, q_1 lie on the same plane. (In fact, they are parallel if P and Q are diametrically opposite or meet at S_1, whose orthogonal projection S over π is the intersection of the lines containing p, q.)

OBSERVATION 2. Let a, b, c be three straight lines in the space such that any two of them are in the same plane. Then all three are in the same plane or they meet at some point, possibly in the infinite, i.e., they may be parallel.

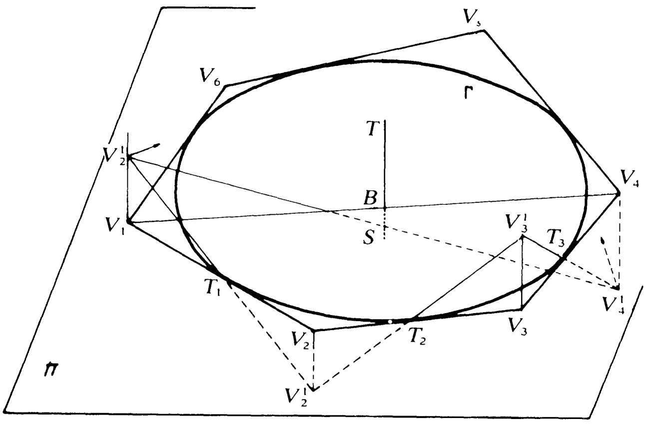


FIG. 1

Consider now the hexagon $V_1V_2V_3V_4V_5V_6$ circumscribed to the circle C (Fig. 1). Call $T_1T_2T_3T_4T_5T_6$ the points of tangency, as indicated. Take an arbitrary point V_1' above the plane π whose orthogonal projection over π is V_1. Join V_1' to T_1 and let V_2' be the intersection of $V_1'T_1$ with the line perpendicular to π through V_2. Join V_2' to T_2 and obtain V_3' in the same way, and so on. Observe that the line $V_6'T_6$ passes through V_1', since all pairs of right triangles formed in the figure are similar.

The lines $V_1'V_2'$, $V_4'V_5'$ are in the situation of the lines determined by p_1, q_1 of the first observation and so the lines $V_1'V_4'$ and $V_2'V_5'$ are in the same plane (the one determined by V_1', V_2', V_4', V_5'). In the same way $V_1'V_4'$ and $V_3'V_6'$ are in a plane and also $V_3'V_6'$ and $V_2'V_5'$. Now, since the three lines $V_1'V_4'$, $V_2'V_5'$, $V_3'V_6'$ are clearly not in a plane (V_1', V_3', V_5' are above π and V_2', V_4', V_6' are below it), according to the second observation, they must meet as some point S. But this implies that their orthogonal projections V_1V_4, V_2V_5, V_3V_6 also meet. So we obtain the theorem of Brianchon. Note that as V_1' varies on the perpendicular V_1W at V_1, point S varies on the perpendicular at the Brianchon point.

Since the configuration leading to S and B in the preceding construction is projective one arrives at the following generalization.

Let $V_1V_2V_3V_4V_5V_6$ be a hexagon whose sides are tangent to a conic C. Let T_1, T_2, T_3, T_4, T_5, T_6, be the points of tangency (Fig. 2). Choose an arbitrary point P on the plane of the hexagon not on the hexagon. Let p_i, $i = 1, 2, \ldots, 6$ be the line joining P to V_i. On p_1 choose an arbitrary point V_1'. Join V_1' to

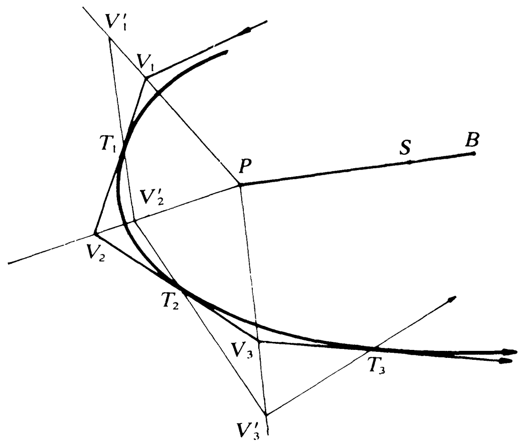


FIG. 2

T_1 and let V_2' be the intersection of this line with p_2. Join V_2' to T_2 and let V_3' be the intersection of $V_2'T_2$ with p_3. And so on. Then $V_1'V_4'$, $V_2'V_5'$, $V_3'V_6'$ meet at a point S. Furthermore, when V_1' moves over p_1, S moves over the line joining P to the point B of Brianchon.

Reference

1. D. Hilbert and S. Cohn-Vossen, Geometry and the Imagination, Chelsea, New York, 1952, p. 105.

FACULTAD DE MATEMÁTICAS, UNIVERSIDAD COMPLUTENSE DE MADRID, MADRID 3, SPAIN.

MATHEMATICAL EDUCATION

EDITED BY PAUL T. MIELKE AND SHIRLEY HILL

Material for this Department should be sent to Paul T. Mielke, Department of Mathematics, Wabash College, Crawfordsville, IN 47933, or to Shirley Hill, Department of Mathematics, University of Missouri, Kansas City, MO 64110.

MATHEMATICS PRE-COLLOQUIUM

S. R. QUINT

1. Introduction. There has been some discussion in these pages ([1], [2], [3]) concerning the desirability of broadening the mathematical education of graduate students. This article describes a program that could be used, toward that general end, in conjunction with some form of a Mathematics Colloquium—here, Colloquium refers to a program of regularly scheduled, mutually independent lectures on current research, given by outside speakers or resident faculty. Based on an existent Colloquium, the Mathematics Pre-Colloquium was designed to provide a built-in series of lectures for graduate students and was instituted at UCLA, at our suggestion, during the winter and spring quarters of 1974.

By its very nature, the Colloquium lecture appeals primarily to those faculty members or graduate students whose research or interest lies near the area of the lecture. The Colloquium has the potential, however, of providing graduate students, in general, with an excellent opportunity for exposure to current research in diverse areas and hence of appealing to a much wider audience. But, unless the student happens to be doing research near the area being discussed, it is very possible that, at best, only the first few minutes of the Colloquium lecture will be understandable. The purpose, then, of the Pre-Colloquium lecture is to make the Colloquium lecture more inviting and more meaningful to those graduate students who have little or no familiarity with the subject of the Colloquium lecture, yet who might be interested in attending while attempting to broaden mathematical background or searching for an area of research.

2. Content of the Pre-Colloquium lecture. The main goal of the Pre-Colloquium lecture is to provide enough background material to make the general ideas of the Colloquium lecture understandable to the interested graduate students. Depending on the Colloquium topic, this material might include definitions, examples, counterexamples and theorems; additionally, if time permits, any of the following would be helpful:

(1) A brief history of the problem.

(2) Relationships, if any, to other areas within or outside of the field of the lecture subject; e.g., if the topic lies in functional analysis, relationships to other areas of analysis, or to algebra, geometry, applied mathematics, or even to disciplines outside of mathematics. Oftentimes, the title of the Colloquium lecture does not provide any clues in this direction and hence might unnecessarily narrow the range of students for whom the Colloquium lecture might be appealing.

(3) Mention of graduate courses in which the student could gain further background for the subject.

(4) Names of faculty members within or outside of the department whose fields of interest are somewhat related to the Colloquium topic. This could provide the student with an opportunity to meet with faculty members for further discussion. (This, in turn, might lead to a thesis advisor.)

(5) A reference list (a dittoed list would save time) of background material which might include: recent relevant research papers; survey articles placing the topic in a more general setting; standard papers, lecture notes or texts.

Our Pre-Colloquium lecture was scheduled for an hour, and was held two days before the Colloquium. The following is a list of some of those lecture topics:

Pre-Colloquium	Colloquium
1. Formalism of Quantum Mechanics	1. The Language of Physics: A Dialect of Mathematics
2. Transcendentality	2. New Applications of Transcendence Theory
3. Importance of the Weil Conjectures	3. Weil Conjectures
4. Introduction to Lax–Phillips Scattering Theory	4. Scattering Theory
5. An Exposition of Carleson's Theorem	5. A New Proof of Carleson's Interpolation Theorem
6. Background Material for the Thursday Colloquium	6. Spectrum of the Laplacian and Lengths of Closed Geodesics on Manifolds

3. Some initial misgivings. Prior to instituting the Pre-Colloquium, the Chairman of the Colloquium Committee had copies of our proposal distributed to the faculty and graduate students, requesting their comments and suggestions. The following were some of their initial misgivings together with the resultant situations after the commencement of the Pre-Colloquium:

(1) Would the apparently most qualified faculty member consent to give the Pre-Colloquium lecture upon merely a few weeks notice?

As soon as the Colloquium speaker/lecturer was scheduled (which was usually more than two weeks prior to the lecture), we requested from the Colloquium Committee Chairman suggestions of potential Pre-Colloquium speakers to be contacted. In all but two cases, the faculty members readily agreed. In one case, the Colloquium speaker was a resident faculty member who felt that his lecture would not require a Pre-Colloquium lecture. (This is one type of situation, not foreseen at the outset, in which a Pre-Colloquium lecture may not be necessary.) In the second case, the faculty member most qualified did not feel sufficiently qualified.

(2) In the case of an outside speaker who provides a "come-on" lecture title, the faculty member nearest to the subject might be in the dark as to the precise material to be dealt with.

In one case, the Pre-Colloquium speaker, considerably in the dark, made a best estimate based on the Colloquium speaker's known publications. Although the Pre-Colloquium turned out to be only tangentially related to the Colloquium, it was nevertheless an interesting and profitable lecture. In another case, the Pre-Colloquium speaker, with a general idea about the probable content of the Colloquium lecture but wanting his lecture to provide a close mesh, contacted the Colloquium speaker to determine what that professor would discuss if he were presenting the Pre-Colloquium.

(3) The Pre-Colloquium lecture might not be appropriate for all Colloquia.

This happened in two cases—one was mentioned in (1) above, and the second was "History of the Differential".

4. Conclusion. The Pre-Colloquium, together with the Colloquium, is one means of providing graduate students with an opportunity to broaden their mathematical backgrounds. It is not clear what level of graduate students would best profit from the lectures. On the one hand, the more advanced students, who have completed the basic courses, would probably be in a better position to understand the lectures and appreciate the possible interrelationships, both of the various branches of mathematics and of mathematics with other disciplines. On the other hand, the less advanced students, who still have some options and flexibility in their courses of study, might obtain some guidance from the lectures.

Acknowledgements. I would like to thank Professor C. C. Moore, UC Berkeley, in particular for a conversation, in 1971, on graduate education and Professors V. S. Varadarajan, P. C. Curtis, Jr., and the faculty of the UCLA mathematics department for their cooperation in this program.

References

1. W. L. Duren, Jr., Are there too many Ph.D.'s in mathematics?, this MONTHLY, 77 (1970) 641–646.
2. I. N. Herstein, On the Ph.D. in mathematics, this MONTHLY, 76 (1969) 818-824.
3. Alex Rosenberg, Some further thoughts on the Ph.D. in mathematics, this MONTHLY, 77 (1970) 646-651.

DEPARTMENT OF MATHEMATICS, OCCIDENTAL COLLEGE, LOS ANGELES, CA 90041.

PROBLEMS AND SOLUTIONS

EDITED BY EMORY P. STARKE

All problems (both elementary and advanced) proposed for inclusion in this Department should be sent to E. P. Starke, 1000 Kensington Ave., Plainfield, NJ 07060. Proposers of problems are urged to enclose any solutions or information that will assist the editors. Ordinarily, problems in well-known textbooks and results in generally accessible sources are not appropriate for this Department. No solutions (except those accompanying proposals) should be sent to Professor Starke.

ELEMENTARY PROBLEMS

Solutions of Elementary Problems should be sent to Problems Group, Faculty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1. To facilitate their consideration, solutions of Elementary Problems in this issue should be typed (with double spacing) and should be mailed before March 31, 1978.

E 2683. *Proposed by Ira Gessel, Massachusetts Institute of Technology*

Let A be the cyclic matrix with $(a_0, a_1, \ldots, a_{p-1})$ as first row, p a prime. If a_i are integers show that $\det A \equiv a_0 + a_1 + \cdots + a_{p-1} \pmod{p}$.

E 2684. *Proposed by Charles A. Nicol, University of South Carolina*

Let A_n be the set of positive integers which are less than n and relatively prime to n. For which n is A_n an arithmetic progression?

E 2685. *Proposed by Ronald Evans, University of California, San Diego, La Jolla*

If p is an odd prime, show that

$$\sum_{i=0}^{p-1} (-1)^i \binom{p^2-p}{pi} \equiv p^{p-1} \pmod{p^p}.$$

E 2686. *Proposed by Peter L. Montgomery, Huntsville, Alabama*

Show that

$$(n+1)\operatorname*{LCM}_{0 \leqq k \leqq n}\left\{\binom{n}{k}\right\} = \mathrm{LCM}\{1, 2, \ldots, n+1\}.$$

E 2687. *Proposed by Ronald Evans, University of California, San Diego, La Jolla*

Does there exist a triangle with rational sides whose base equals its altitude?

E 2688. *Proposed by David Jackson, University of Waterloo, Ontario, Canada*

Let $\{f_i\}$ and $\{g_i\}$ $(i = 0, 1, 2, \ldots)$ be the solutions of the recurrence equation

$$u_{m+1} = -u_m - m(m+1)xu_{m-1}$$

satisfying the initial conditions $f_0 = 0$, $f_1 = 1$ and $g_0 = 1$, $g_1 = -1$, respectively.

Show that the coefficient of x^{n-1} in the Maclaurin expansion of $-f_n/g_n$ is t_{2n-1} where

$$\tan x = \sum_{n \geqq 1} t_{2n-1} \frac{x^{2n-1}}{(2n-1)!}.$$

SOLUTIONS OF ELEMENTARY PROBLEMS

A Formula for a Function

E 2604 [1976, 483]. *Proposed by E. T. H. Wang, Wilfrid Laurier University, Waterloo, Ontario, Canada*

Let $\mathbf{N} = \{0, 1, 2, \ldots\}$ and let $A : \mathbf{N} \to \mathbf{N}$ be defined by $A(n) = [2n/3]$. *For* $n \in \mathbf{N}$ let $k \in \mathbf{N}$ be the smallest integer such that $A^k(n) = 0$ and define $f(n) = k$. Find a formula, as simple as possible, for the function f.

Solution by Phillip C. Washburn, Hyattsville, Maryland (revised by the editor). Since A is monotonic it follows that f is also monotonic. Since f maps $\mathbf{N}$ onto $\mathbf{N}$, there exists a strictly increasing sequence $s_0, s_1, s_2, \ldots$ such that

$$f^{-1}(n) = \{m \in \mathbf{N} \mid s_{n-1} < m \leqq s_n\}$$

holds for $n \in \mathbf{N}$ $(s_{-1} = -1)$. For instance, $s_0 = 0$, $s_1 = 1$, $s_2 = 2$, $s_3 = 4$, $s_4 = 7$.

The term s_{n+1} is the largest integer m such that $A(m) = s_n$. It follows that

$$s_{n+1} = \left[\frac{3}{2} s_n\right] + 1 \qquad (n \geqq 0). \tag{1}$$

We claim that s_n is odd for infinitely many values of n. Otherwise there is an m such that s_n is even for $n \geqq m$. This gives

$$s_{m+k} = \left(\frac{3}{2}\right)^k s_m + \sum_{i=0}^{k-1} \left(\frac{3}{2}\right)^i = \left(\frac{3}{2}\right)^k (s_m + 2) - 2.$$

However, for large k the last expression is not an integer.

From (1) we have $\frac{1}{2} \leqq s_{n+1} - \frac{3}{2} s_n \leqq 1$. Hence if $u_n = (\frac{2}{3})^n s_n$ then

$$\frac{1}{3}\left(\frac{2}{3}\right)^n \leqq u_{n+1} - u_n \leqq \left(\frac{2}{3}\right)^{n+1}.$$

It follows that $\Sigma_{n \geqq 0}(u_{n+1} - u_n) = \lim_{n \to \infty} u_n = c$ exists and that

$$\frac{1}{3} \sum_{k \geqq n} \left(\frac{2}{3}\right)^k \leqq c - u_n < \sum_{k \geqq n+1} \left(\frac{2}{3}\right)^k. \tag{2}$$

The last inequality sign is strict because s_n is odd for infinitely many values of n. We can rewrite (2) as

$$\left(\frac{2}{3}\right)^n \leqq c - u_n < 2\left(\frac{2}{3}\right)^n,$$

$$1 \leqq \left(\frac{3}{2}\right)^n c - s_n < 2,$$

and so

$$s_n = \left[\left(\frac{3}{2}\right)^n c\right] - 1.$$

From $s_{f(n)-1} < n \leqq s_{f(n)}$ we obtain successively

$$\left[\left(\frac{3}{2}\right)^{f(n)-1} c\right] < n+1 \leqq \left[\left(\frac{3}{2}\right)^{f(n)} c\right],$$

$$\left(\frac{3}{2}\right)^{f(n)-1} c < n+1 \leqq \left(\frac{3}{2}\right)^{f(n)} c,$$

$$f(n)-1 < \log_{3/2}\left(\frac{n+1}{c}\right) \leqq f(n),$$

$$f(n) = -\left[-\log_{3/2}\left(\frac{n+1}{c}\right)\right].$$

Also solved by R. W. K. Odoni (England), David Penney, David Stone, James Walker and Abraham Ziv (Israel).

Editor's Comment. Washburn gives the approximate value $c \doteqdot 1.622705$. It may be interesting to know whether c is irrational or even transcendental.

Labels on a Chessboard

E 2605 [1976, 566]. *Proposed by Andreas P. Hadjipolakis, Anopolis Sfakion, Crete*

Consider a chessboard of odd order n ($n \geqq 5$). Assign label m to a cell of the chessboard if it can be reached by the knight in m steps starting from the central cell and this m is minimal. Determine the number $K(n; m)$ of cells labelled m.

Solution by Roger Weitzenkamp, Oak Park, Illinois. We consider first the case of an infinite chessboard. Let $L \subset \mathbf{Z}^2$ be the lattice formed by the centers of the squares of the chessboard. We assume that $(0,0)$ is the center of the initial square (with label 0). Denote by L_m the set of centers of the squares with label m. We say that $(a, b) \in L$ is even or odd according to whether $a+b$ is even or odd.

Let O_m be the closed octagon with vertices $(\pm m, \pm 2m), (\pm 2m, \pm m)$ with all choices of the signs. Let A_m be the half-open annulus $O_m \backslash O_{m-2}$ ($m \geqq 3$).

PROPOSITION 1. *For $m \geqq 5$, L_m is the set of lattice points in A_m which have the same parity as m.*

Proof. We use induction on m. The claim may be verified by inspection when $m=5$ or $m=6$. Assume it is true for all m such that $5 \leqq m \leqq k$ where $k \geqq 6$. All points in $L \cap O_{k-1}$ have labels $\leqq k$ and $L_i \subset O_k$ for $i \leqq k$. Therefore $L_{k+1} \subset A_{k+1}$. Since each knight's move connects lattice points of different parities it is clear that each point in L_{k+1} has the same parity as $k+1$. It remains to verify that every lattice point in A_{k+1} having the same parity as $k+1$ is a knight's move from a point in L_k. Since L_k is known, this is not difficult and we leave it to the reader.

Let S_n be the closed square with vertices $(\pm n, \pm n)$. It is easy to see that if $(a, b) \in L_m \cap S_n$ ($m \geqq 5$) then one can reach (a, b) from $(0,0)$ in m knight's moves without leaving S_n. Therefore we have $K(n; m) = |S_n \cap L_m|$ for $m \geqq 5$.

PROPOSITION 2. *For $m \geqq 5$ and odd n we have*

$K(n;m) = 0$	for $n < 3(m-1)$	
$= 4$	for $n = 3m-3$	(m even)

$$\begin{aligned}
&= 16 && \text{for } n = 3m - 1 && (m \text{ even})\\
&= 8 && \text{for } n = 3m - 2 && (m \text{ odd})\\
&= 12(n - 3m + 2) && \text{for } 3m \leqq n \leqq 4m - 7 &&\\
&= 16m - 44 && \text{for } n = 4m - 5 &&\\
&= 20m - 32 && \text{for } n = 4m - 3 &&\\
&= 24m - 24 && \text{for } n = 4m - 1 &&\\
&= 28m - 20 && \text{for } n \geqq 4m + 1\,. &&
\end{aligned}$$

Sketch of Proof for odd $m \geqq 7$. If $n < 3(m-1)$ then $S_n \cap A_m = \varnothing$ and so $K(n;m) = 0$. When n increases we get first overlap of S_n and A_m for $n = 3m - 2$ and then $|S_n \cap L_m| = 8$. In the range $3m \leqq n \leqq 4m - 7$ we are moving along the slanted sides of the octagonal annulus and are picking up 24 points of L_m for each increase in n. Here $K(n;m) = 24 + 24(n - 3m)/2 = 12(n - 3m + 2)$. When $4m - 5 \leqq n \leqq 4m + 1$, we pick up $4(m+4)$, $4(m+3)$, $4(m+2)$, and $4(m+1)$ points of L_m with each increase in n, giving

$$K(4m-5;m) = 12(4m - 7 - 3m + 2) + 4(m+4) = 16m - 44,$$

$$K(4m-3;m) = 16m - 44 + 4(m+3) = 20m - 32,$$

$$K(4m-1;m) = 24m - 24,$$

$$K(4m+1;m) = 28m - 20.$$

Thereafter $K(n;m)$ does not increase with n ($L_m \subset S_n$), so that $K(n;m) = 28m - 20$ for $n \geqq 4m + 1$.

For $m \leqq 4$, the results are given in the table below

	$K(n;m)$			$m \leqq 4$
$n \backslash m$	1	2	3	4
5	8	8	4	4
7	8	20	16	4
9	8	32	32	8
11	8	32	52	28
13	8	32	68	52
15	8	32	68	76
$\geqq 17$	8	32	68	96

Also solved by John Beidler, Jordi Dou (Spain), Roger Eggleton (Australia), Eli Isaacson, James Rue, Gillian Valk and Wayne Wild. Partial solutions by Daniel Cohen, and Roger Lyndon.

Pell's Equation in Disguise

E 2606 [1976, 566]. *Proposed by R. S. Luthar, University of Wisconsin, Janesville*

Show that there are infinitely many integers n such that $2n + 1$ and $3n + 1$ are both perfect squares, and that such n must be multiples of 40.

Solution assembled from many submissions. Eliminating n from the equations

$$2n + 1 = x^2, \qquad 3n + 1 = y^2 \tag{1}$$

yields

$$3x^2 - 2y^2 = 1. \tag{2}$$

The unimodular transformation $x = u + 2v$, $y = u + 3v$ transforms (2) into Pell's equation

$$u^2 - 6v^2 = 1. \tag{3}$$

From the theory of Pell's equation it is known that all non-negative integral solutions (u_k, v_k) of (3) are given by

$$u_k + v_k\sqrt{6} = (5 + 2\sqrt{6})^k, \qquad k = 0, 1, 2, \ldots .$$

Further, any solution (n, x, y) of (1) must have x odd, thus n even, and hence y odd, showing that $8|n$. Finally, (1) implies that $x^2 + y^2 \equiv 2 \pmod 5$, thus $x^2 \equiv y^2 \equiv 1 \pmod 5$, and $5|n$.

Solved by 75 readers and the proposer.

A Functional Equation

E 2607 [1976, 566]. *Proposed by the Eastern Montana College Problem Group*

Solve the functional equation

$$f(x, y) + f(y, z) + f(z, x) = 3f(\tfrac{1}{3}(x + y + z), \tfrac{1}{3}(x + y + z))$$

in the class of all continuous functions $\mathbb{R}^2 \to \mathbb{R}$.

What can be said about the solutions in the class of all functions $\mathbb{R}^2 \to \mathbb{R}$?

Solution by John A. Baker, University of Waterloo, Canada. We solve the equation

$$f(x, y) + f(y, z) + f(z, x) = F(x + y + z), \tag{1}$$

where $f: K^2 \to L$ and $F: K \to L$, K and L being additive abelian groups and division by 2 being uniquely defined in L, that is, for every a in L there is a unique x, denoted by $\frac{1}{2}a$, in L such that $2x = a$. (Note that in the case where K, like $\mathbb{R}$, admits unique division by 3, $F(x)$ necessarily equals $3f(\frac{1}{3}x, \frac{1}{3}x)$.)

Suppose that (1) holds and let $g(x, y) = \frac{1}{2}(f(x, y) - f(y, x))$, $h(x, y) = \frac{1}{2}(f(x, y) + f(y, x)) - \gamma$ where $\gamma = f(0, 0)$; thus $f(x, y) = g(x, y) + h(x, y) + \gamma$. Interchanging x and y in (1) and subtracting the resulting equation from (1) gives (after division by 2)

$$g(x, y) + g(y, z) + g(z, x) = 0. \tag{2}$$

Similarly, adding the resulting equation to (1) gives

$$h(x, y) + h(y, z) + h(z, x) = G(x + y + z) \tag{3}$$

where $G(x) = F(x) - 3\gamma$.

To solve for $g(x, y)$ put $\phi(x) = g(x, 0)$. Then $g(0, x) = -\phi(x)$ since $g(x, y) = -g(y, x)$, and $z = 0$ in (2) gives $g(x, y) = \phi(x) - \phi(y)$.

To find $h(x, y)$, put $\psi(x) = h(x, 0)$. Then $h(0, x) = \psi(x)$ since $h(x, y) = h(y, x)$. Also $h(0, 0) = 0$ and thus $y = z = 0$ in (3) gives $G(x) = 2\psi(x)$. Using this in (3) with $z = 0$, we obtain $h(x, y) = 2\psi(x + y) - \psi(x) - \psi(y)$.

We can now express (3) in terms of ψ:

$$\psi(x + y) + \psi(y + z) + \psi(z + x) - \psi(x) - \psi(y) - \psi(z) = \psi(x + y + z). \tag{4}$$

Let $a(x) = \frac{1}{2}(\psi(x) - \psi(-x))$ and $q(x) = \frac{1}{2}(\psi(x) + \psi(-x))$, so that $\psi(x) = q(x) + a(x)$. Since $\psi(0) = 0$, (4) with $z = -x - y$ becomes $a(x + y) = a(x) + a(y)$, and a is additive. Also $z = -y$ in (4) gives $\psi(x + y) + \psi(x - y) = 2\psi(x) + \psi(y) + \psi(-y)$ which, added to the equation obtained from it on replacing x by $-x$ and y by $-y$, gives $q(x + y) + q(x - y) = 2q(x) + 2q(y)$, that is, q is (homogeneous) quadratic.

Combining the above and using the additivity of a, we find that

$$f(x,y)=\phi(x)-\phi(y)+2q(x+y)-q(x)-q(y)+a(x)+a(y)+\gamma$$

where a is additive and q is quadratic. Conversely, it is easily checked that every such f, with ϕ and γ an arbitrary function and constant respectively, provides a solution of (1).

An alternative form of the solution can be given using symmetric biadditive functions in place of quadratic functions. Define $b(x,y)=\frac{1}{2}(q(x+y)-q(x)-q(y))$. Then $b(x,y)=b(y,x)$ and the quadratic identity for q implies fairly readily that $b(x+y,z)=b(x,z)+b(y,z)$, so that b is biadditive; moreover, $q(x)=b(x,x)$. It may be verified, and it is known (J. Aczél, *The general solution of two functional equations by reduction to functions additive in two variables and with the aid of Hamel bases*, Glasnik Mat.-Fiz. Astr. 20 (1965), 65–72), that this correspondence between quadratic and symmetric biadditive functions is one-one (given that L admits unique division by 2). It follows that the general solution of (1) may be written as

$$f(x,y)=\phi(x)-\phi(y)+b(x,x)+4b(x,y)+b(y,y)+a(x)+a(y)+\gamma$$

where ϕ, a, γ are as above and b is symmetric biadditive.

For $K=L=\mathbb{R}$, the most general additive and biadditive functions can be found using a Hamel basis. For continuous solutions, note that the continuity of f is equivalent to that of ϕ, a, and b. Now it is well known that the only continuous additive functions $a:\mathbb{R}\to\mathbb{R}$ are of the form $a(x)=\alpha x$; similarly the only continuous biadditive functions $b:\mathbb{R}^2\to\mathbb{R}$ are of the form $b(x,y)=\beta xy$. Thus the continuous solutions of (1) are given by

$$f(x,y)=\phi(x)-\phi(y)+\beta(x^2+4xy+y^2)+\alpha(x+y)+\gamma,$$

where ϕ is a continuous but otherwise arbitrary function and α, β, γ are arbitrary constants.

Also solved by Detlef Laugwitz (Germany), O. P. Lossers (Netherlands), M. A. McKiernan (Canada), and, in the continuous case, Robert Breusch, and Christopher Henley.

Editor's comment. McKiernan remarks that the solution must be continuous (and hence as given above) if it is known to be measurable, or continuous at one point, etc. In connection with the correspondence between quadratic and biadditive functions he cites S. Mazur and W. Orlicz, *Grundlegende Eigenschaften der Polynomischen Operationen*, Studia Math. 5 (1934), 50–68 and 179–189 (as also does the paper by Aczél referred to by Baker).

A Difference Equation in Two Variables

E 2609 [1976, 567]. *Proposed by Glen E. Bredon, Rutgers University*

Define integers a_{ij} $(i,j\geqq 1)$ by $a_{i1}=a_{1j}=1$ and

$$a_{ij}=i\,a_{i,j-1}+j\,a_{i-1,j}\qquad (i,j\geqq 2).$$

Show that

$$\sum_{i=1}^{2n-1}(-1)^{i-1}a_{i,2n-i}\equiv 1\qquad (\mathrm{mod}\,3).$$

I. *Solution by Aage Bondesen, Espergaerde, Denmark* (*revised by the editor*). Let $a_{ij}=0$ if $i+j\geqq 2$ and $i\leqq 0$ or $j\leqq 0$. Then we have $a_{ij}=i\,a_{i,j-1}+j\,a_{i-1,j}$ for all (i,j) such that $i+j\geqq 3$. Let

$$S_n=\sum_{i=1}^{n}(-1)^{i-1}a_{i,n+1-i}=\sum_{-\infty}^{+\infty}(-1)^{i-1}a_{i,n+1-i}.$$

Since

$$a_{ij}=i\,a_{i,j-1}+j\,a_{i-1,j}=i^2a_{i,j-2}+(2ij-i-j)a_{i-1,j-1}+j^2a_{i-2,j}$$

$$= i^3 a_{i,j-3} + [3i(ij-i-j)+i+j]a_{i-1,j-2}$$
$$+ [3j(ij-i-j)+i+j]a_{i-2,j-1} + j^3 a_{i-3,j},$$

we have

$$a_{ij} \equiv i a_{i,j-3} + (i+j)(a_{i-1,j-2} + a_{i-2,j-1}) + j a_{i-3,j} \pmod 3.$$

Thus for $n \geqq 4$ we have

$$S_n \equiv \sum (-1)^{i-1}[i a_{i,n-2-i} + (n+1)(a_{i-1,n-1-i} + a_{i-2,n-i}) + (n+1-i)a_{i-3,n+1-i}]$$

$$= \sum (-1)^{i-1} a_{i,n-2-i}[i-(n+1)+(n+1)-(n+1-i-3)]$$

$$= \sum (-1)^{i-1}(2i-n+2)a_{i,n-2-i}.$$

On the other hand

$$S_{n-2} = \sum (-1)^{i-1} a_{i,n-1-i} = \sum (-1)^{i-1}[i a_{i,n-2-i} + (n-1-i)a_{i-1,n-1-i}]$$

$$= \sum (-1)^{i-1} a_{i,n-2-i}[i-(n-1-i-1)] = \sum (-1)^{i-1}(2i-n+2)a_{i,n-2-i}.$$

Thus $S_n \equiv S_{n-2} \pmod 3$ for $n \geqq 4$ but it is also true for $n = 3$. Hence $S_{2n-1} \equiv S_1 = 1 \pmod 3$. Note also that $S_{2n} = 0$.

II. *Generalization by Allan Wm. Johnson, Jr., Defense Communications Agency, Washington, D.C.* The Eulerian numbers

$$A(n,k) = \sum_{m=0}^{k} (-1)^m \binom{n+1}{m}(k-m)^n \qquad (1 \leqq k \leqq n)$$

satisfy $A(n,k) = A(n, n-k+1)$ and

$$A(n+1,k) = (n-k+2)A(n,k-1) + kA(n,k) \qquad (2 \leqq k \leqq n).$$

For this we refer to L. Comtet, *Analyse Combinatoire*, Paris 1970, vol. 1, 63–64 and vol. 2, 82–86. Since $A(n,1) = 1$ it follows that $A(n,n) = 1$. Therefore we have

$$a_{ij} = A(i+j-1, j) = \sum_{k=0}^{i-1} (-1)^k \binom{i+j}{k}(i-k)^{i+j-1}.$$

Thus

$$S_n = \sum_{i=1}^{n} \sum_{k=0}^{i-1} (-1)^{k-i+1}\binom{n+1}{k}(i-k)^n$$
$$= \sum_{k=0}^{n-1} \sum_{i=k+1}^{n} (-1)^{k-i+1}\binom{n+1}{k}(i-k)^n$$
$$= \sum_{k=0}^{n-1} \binom{n+1}{k} \sum_{j=1}^{n-k} (-1)^{j-1} j^n.$$

Let E and Δ be the shift and difference operator, respectively. Thus $(Ef)(x) = f(x+1)$ and $(\Delta f)(x) = f(x+1) - f(x)$ for any function f. We find that

$$S_n = \sum_{k=0}^{n-1} \binom{n+1}{k} \left(\sum_{j=1}^{n-k} (-1)^{j-1} E^j \right) (x^n) \Big|_{x=0}$$

$$= (1+E)^{-1} \sum_{k=0}^{n-1} \binom{n+1}{k} (E + (-E)^{n-k+1})(x^n) \Big|_{x=0}$$

$$= (1+E)^{-1}[(2^{n+1}-1)(1+E) - 2^{n+1} + (1-E)^{n+1}](x^n) \Big|_{x=0}.$$

Since $(1-E)^{n+1}(x^n) = 0$ and $1 + E = 2 + \Delta$ we obtain

$$S_n = -2^n \left(1 + \frac{\Delta}{2}\right)^{-1} (x^n) \Big|_{x=0},$$

(1) $$S_n = \sum_{i=1}^{\infty} (-1)^{i-1} 2^{n-i} \Delta^i 0^n$$

where $\Delta^i 0^n = \Delta^i (x^n) \Big|_{x=0}$.

Now we can prove the following proposition:

THEOREM. *If p is an odd prime and $m \equiv n \pmod{p-1}$ then $S_m \equiv S_n \pmod{p}$.*

Proof. Since $i^m \equiv i^n \pmod{p}$ by Fermat's theorem and

$$\Delta^i 0^n = \sum_{k=0}^{i} (-1)^k \binom{i}{k} (i-k)^n$$

we see that $\Delta^i 0^m \equiv \Delta^i 0^n \pmod{p}$. The claim now follows from (1).

Also solved by David Anderson, John Bailar, III, M. T. Bird, D. M. Bloom, Peter de Buda, L. Carlitz, Eli Isaacson, Lael Kinch, L. E. Mattics, Adam Riese, Robert Sloan, Gillian Valk, and G. Wedderburn.

Comment. Several solvers observe that

$$S_{n-1} = \frac{1}{n} 2^n (2^n - 1) B_n \qquad (n \geqq 2),$$

where B_n are Bernoulli numbers. Also we have

$$\tan x = \sum_{n \geqq 1} (-1)^{n-1} S_{2n-1} \cdot \frac{x^{2n-1}}{(2n-1)!}.$$

Johnson gives the reference to Uspensky and Heaslet, *Elementary Number Theory*, McGraw Hill, 1939, Ex. 4, on pp. 267–268 where the above proposition is stated as an exercise.

An Impossible Partition

E 2613 [1976, 656]. *Proposed by D. E. Knuth, Stanford University, and the Mayaguez Problems Group, University of Puerto Rico (independently).*

Partition the real line $\mathbb{R}$ into a countable union of compact subsets.

Solution by Bert Gunter, Beloit College, Wisconsin, and O. P. Lossers, Technological University, Eindhoven, Netherlands (independently). We prove that such partition does not exist. In fact the following stronger result will be proved: *It is impossible to partition $\mathbb{R}$ into countably many closed subsets $C(i)$, $0 \leqq i < \infty$.*

Proof. Choose $x, y \in \mathbb{R}$ so that $x \in C(0)$ and $y \notin C(0)$. Let, say, $x < y$ and define $a_0 =$

$\sup(C(0)\cap(-\infty,y))$. Then $a_0\in C(0)$ and $C(0)\cap(a_0,y)=\varnothing$. Let i_1 be the smallest positive integer such that $C(i_1)\cap(a_0,y)\neq\varnothing$, and put $b_0=\inf(C(i_1)\cap(a_0,y))$. Then $b_0\in C(i_1)$, $a_0<b_0$, and the sets $C(i)$ for $i\leqq i_1$ do not meet (a_0,b_0). Now let i_2 be the smallest positive integer such that $C(i_2)\cap(a_0,b_0)\neq\varnothing$ and put $a_1=\sup(C(i_2)\cap(a_0,b_0))$. We have $a_1\in C(i_2)$, $a_0<a_1<b_0$, and the sets $C(i)$ for $i\leqq i_2$ do not meet (a_1,b_0). Next we let i_3 be the smallest positive integer such that $C(i_3)\cap(a_1,b_0)\neq\varnothing$ and put $b_1=\inf(C(i_3)\cap(a_1,b_0))$. Then $b_1\in C(i_3)$, $a_0<a_1<b_1<b_0$, and the sets $C(i)$ for $i\leqq i_3$ do not meet (a_1,b_1). This construction can be continued *ad infinitum* by alternating sup and inf. The sequence of segments $[a_i,b_i]$ is nested and hence there exists z such that $z\in[a_i,b_i]$ for all $i\geqq 0$. But then it follows from our construction that $z\notin C(i)$ for all $i\geqq 0$. This is a contradiction.

Also solved by John Baker (Canada), Charles Blair, Eric Brosius, Melvin Henriksen, R. Hodel, Edward Howorka, Burton Jones, J. Lawlor, James Mauldon, Mark Meyerson, Stephen Noltie, Edward Ordman, Nicholas Passell, Horace Smith, Arthur Solomon, Albert Wilansky, and the proposers.

Editor's Comments. Baker, Lawlor and Ordman note that this problem is essentially equivalent to Exercise 10.2, p. 62 of R. P. Boas, *A Primer of Real Functions*; solution appears on p. 178.

Henriksen, Hodel, Jones, and Wilansky note that the result proved above follows from a theorem of W. Sierpiński which states that a continuum has no countable partition into closed subsets. One has just to apply this theorem to the one-point compactification of $\mathbb{R}$. For this theorem see K. Kuratowski, *Topology*, vol. II, p. 173. Wilansky also refers to the papers: C. Eberhart, J. B. Fugate, L. Mohler, *Spaces which cannot be written as a countable disjoint union of closed subsets*, Canad. Math. Bull. 16(1973), 435–437, and G. J. O. Jameson, *A variant of a theorem of Sierpiński concerning partitions of continua*, Colloquium Math. 25(1972), 79–80, for further generalizations.

ADVANCED PROBLEMS

All solutions of Advanced Problems should be sent to J. Barlaz, Hill Center, Rutgers University, New Brunswick, N. J. 08903. Solutions of Advanced Problems in this issue should be typed (with double spacing) on separate signed sheets and should be mailed before March 31, 1978.

An asterisk (*) means neither the proposer nor the editors supplied a solution.

6180. *Proposed by L.C. Larson, Saint Olaf College, Minnesota*

Let A and B be ideals of a commutative ring R with unity. Show that $\{x\in R\colon xB\subseteq xA\}$ is an ideal if R is either an integral domain or a principal ideal ring, but that in general it need not be.

6181*. *Proposed by J.M. Arnaudies, Strasbourg, France*

Let n be an integer $\geqq 3$, and $A_0, A_1,\ldots,A_n$ be n single-valued real functions defined and continuous on a given topological Hausdorff space T. Suppose that for all $t\in T$, the 2-form

$$A_0x^n + A_1x^{n-1}y + \cdots + A_ny^n$$

(where the A_i take their values for t) defines n real distinct lines in the 2-dimensional real projective space.

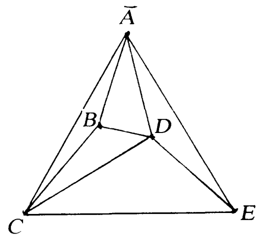
Characterize spaces T such that, for any choice of the A_i, there exists a system of continuous functions $(P_1, Q_1, P_2, Q_2,\ldots,P_n, Q_n)$, real-valued, defined on T, satisfying the formal equality,

$$A_0x^n + A_1x^{n-1}y + \cdots + A_ny^n = (P_1x+Q_1y)(P_2x+Q_2y)\cdots(P_nx+Q_ny).$$

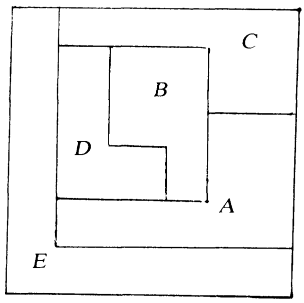
6182*. *Proposed by A.K. Austin, University of Sheffield, England*

Prove or disprove that any finite planar graph can be represented by a map in which all the regions

are L-shaped with sides horizontal and vertical. For example



can be represented by



6183*. *Proposed by Albert A. Mullin, Ft. Hood, Texas*

Let R be a ring with a finite number n of multiplicative identities.
(a) If R is commutative, show that n is a power of 2.
(b) If R has a unit, show that n is even but need not be a power of 2.
(c) Is there an R for which n is an odd prime?

6184. *Proposed by Ole Jorsboe, Technical University of Denmark*

Let $(\phi_n)_{n=1}^{\infty}$ be an orthonormal system of real-valued piecewise continuous functions on the interval $[0,1]$ with the property that if f is a real-valued piecewise continuous function on $[0,1]$ fulfilling $(f,\phi_n)=\int_0^1 f(x)\phi_n(x)dx=0$ for all $n\in N$, then f is 0 at all points of continuity.

Does this imply that (ϕ_n) spans the space of all real-valued piecewise continuous functions on $[0,1]$, i.e., can every piecewise continuous function f be written in the form

$$f=\lim_{N\to\infty}\sum_{n=1}^{N}a_n\phi_n\ ?$$

6185. *Proposed by John Milcetich, Federal City College*

Let $f(z,\theta)=(1+e^{i\theta}z)^{\beta}(1-z)^{-\alpha}$ where $|z|<1$, $\theta\in R$ and $\alpha\geqq\beta\geqq 1$. Show that for $p>0$ and $0<r<1$

$$\int_{-\pi}^{\pi}|f'(re^{i\phi},\theta)|^p d\phi\leqq\int_{-\pi}^{\pi}|f'(re^{i\phi},0)|^p d\phi.$$

SOLUTIONS OF ADVANCED PROBLEMS

Linear Functionals in Normed Spaces

6078 [1976, 205]. *Proposed by Albert Wilansky, Lehigh University*

Is it possible for a continuous linear functional on a normed space to map every bounded closed set onto a closed set of scalars? (Exclude trivial cases.)

I. *Solution by Seymour Goldberg, University of Maryland.* For every non-zero continuous linear functional f on an infinite dimensional normed linear space X, there exists a countable, bounded closed set $S\subset X$ such that $f(S)$ is not closed.

Since the kernel N of f is infinite dimensional, there exists a sequence $\{x_n\}$ in the 1-sphere of N such that $\{x_n\}$ has no convergent subsequence. For $v \notin N$, $S = \{x_n + (1/n)v : n = 1, 2, \ldots\}$ is closed and bounded, yet $f(S) = \{(1/n)f(v)\}$ is not closed.

REMARK. If T is a closed linear operator with domain and range in a Banach space and the kernel of T is finite dimensional, then T maps bounded closed sets onto closed sets if and only if T has a closed range. This is theorem IV.1.10, in Seymour Goldberg, *Unbounded Linear Operators*, McGraw-Hill 1966. The above example shows that the assumption that the kernel be finite dimensional is essential.

II. *Note by Robert C. James, Claremont Graduate School.* The answer is quite different if "bounded closed set" is replaced by "bounded closed convex set." Such a set is weakly closed and therefore weakly compact if the space is reflexive, in which case a continuous linear functional (being weakly continuous) maps the set onto a compact set. However, if a space is not reflexive, then there is a continuous linear functional which does not attain its supremum on the unit ball and therefore maps the closed unit ball onto an open interval [M. M. Day, *Normed Linear Spaces*, Springer-Verlag, New York 1973, pp. 63–64].

Also solved by J. Borwein & J. Phillips (Canada), Maxim Enders (South Africa), Detlef Laugwitz (Germany), Geoffrey Robinson (England), Alan Shuchat, Don Swartwout, Aaron Todd, and the proposer.

Nowhere Continuous, Quasi-continuous Functions

6081 [1976, 205]. *Proposed by T. Šalát, University of J. A. Komensky, Czechoslovakia*

Let (X, d) be a metric space. We call $f: X \to R$ quasi-continuous at x_0 if for each positive ε and δ there exists an open sphere $S(x_1, \delta_1) = \{x : d(x, x_1) < \delta_1\} \subset S(x_0, \delta)$, such that $f[S(x_1, \delta_1)] \subset (f(x_0) - \varepsilon, f(x_0) + \varepsilon)$. Does there exist a metric space of first category and with no isolated points which allows a quasi-continuous function which is nowhere continuous?

Solution by James L. Cornette, Iowa State University. A well-known real variable function with domain restricted to the rationals provides the answer. Let Q denote the rational numbers, let $\phi: Q \to \{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots\}$ be a one to one map, and let $f: Q \to R$ be defined by

$$f(x) = \sum_{r \leq x} \phi(r).$$

Then f is continuous from the right at every rational and is therefore quasi-continuous, but f is discontinuous from the left at every rational, and is therefore nowhere continuous.

Also solved by William Gorman III, Togo Nishimura, Eric van Douwen, and Don Swartwout.

Nishimura's example makes use of a function which is found in S. Kempisty, *Sur les fonctions quasicontinues*, Fund. Math. 19 (1932), pp. 184–197. Van Douwen and Swartwout note that the condition in the problem "first category without isolated points" is redundant. Van Douwen constructs a bijection $f: Q \to Q$ such that both f and f^{-1} are quasi-continuous but nowhere continuous.

The Functions $\Sigma_r p / r(p + r)$, $\Sigma_r (-1)^{r-1} \binom{p}{r} / r$

6083 [1976, 205]. *Proposed by Emil Grosswald, Temple University*

Prove that, for real $p > 0$, the following identity holds:

$$\sum_{r=1}^{\infty} \frac{p}{r(p+r)} = \sum_{r=1}^{\infty} (-1)^{r-1} \binom{p}{r} \frac{1}{r}. \tag{1}$$

What is the function represented by both sides of this identity?

Solution by L. Van Hamme, Vrije Universiteit Brussel, Belgium. It is known that

$$\sum_{r=1}^{\infty} \frac{p}{r(p+r)} = \sum_{r=1}^{\infty} \left(\frac{1}{r} - \frac{1}{r+p}\right) = \psi(p) + \gamma + \frac{1}{p},$$

where $\psi(p)=(d/dp)\log\Gamma(p)$ and γ is the Euler constant. In Milne-Thomson, *The Calculus of Finite Differences* (Macmillan, 1933) on p. 315, one finds

$$\psi(x)+\gamma=\binom{x-1}{1}-\frac{1}{2}\binom{x-1}{2}+\frac{1}{3}\binom{x-1}{3}-\cdots. \tag{2}$$

Replacing x by $p+1$ one obtains the desired formula.

Formula (2) is valid for all complex numbers x with $\mathrm{Re}(x)>0$ and hence (1) is true for all complex p with $\mathrm{Re}(p)>-1$.

We can get the more general result:

$$\sum_{r=0}^{\infty}\frac{p}{(x+r)(x+p+r)}=\sum_{s=1}^{\infty}\frac{(-1)^{s-1}}{s}\cdot\frac{p(p-1)\cdots(p-s+1)}{x(x+1)\cdots(x+s-1)},\ \mathrm{Re}(x+p)>0, \tag{3}$$

and (1) follows upon putting $x=1$. To prove (3) we start with the formula

$$\frac{1}{x+p}-\frac{1}{x}=\sum_{s=1}^{\infty}(-1)^s\frac{p(p-1)\cdots(p-s+1)}{x(x+1)\cdots(x+s)},\quad \mathrm{Re}(x+p)>0. \tag{4}$$

This result is well known and can be proved in many ways (see e.g. the book of Milne-Thomson cited above).

It is also a special case of the following formula for the hypergeometric function

$$F(a,b;c;1)=\frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)},\quad \mathrm{Re}(c-a-b)>0$$

as can be seen by putting $a=-p$, $b=1$, $c=x+1$.

If we replace x by $x+r$ in (4) and sum for $r=0,1,\dots,R$ we obtain

$$\sum_{r=0}^{R}\left(\frac{1}{x+r+p}-\frac{1}{x+r}\right)=\sum_{s=1}^{\infty}(-1)^s p(p-1)\cdots(p-s+1)\sum_{r=0}^{R}\frac{1}{(x+r)(x+r+1)\cdots(x+r+s)}.$$

Since

$$\begin{aligned}&\sum_{r=0}^{R}\frac{1}{(x+r)\cdots(x+r+s)}\\ &=\sum_{r=0}^{R}\frac{1}{s}\left[\frac{1}{(x+r)(x+r+1)\cdots(x+r+s-1)}-\frac{1}{(x+r+1)\cdots(x+r+s)}\right]\\ &=\frac{1}{s\cdot x(x+1)\cdots(x+s-1)}-\frac{1}{s(x+R+1)\cdots(x+R+s)},\end{aligned}$$

we have

$$\begin{aligned}&\sum_{r=0}^{R}\frac{p}{(x+r)(x+p+r)}\\ &=\sum_{s=1}^{\infty}(-1)^{s-1}\frac{p(p-1)\cdots(p-s+1)}{s\cdot x(x+1)\cdots(x+s-1)}-\sum_{s=1}^{\infty}(-1)^{s-1}\frac{p(p-1)\cdots(p-s+1)}{s(x+R+1)\cdots(x+R+s)}.\end{aligned}$$

It follows from the theory of series of inverse factorials that the last series is uniformly convergent, with respect to R, for R sufficiently great. Hence if R tends to infinity we may calculate the limit termwise to complete the proof.

Also solved by the proposer and fifty-one other readers.

Notes. (1) R. Tremblay and M. L. Lavertu, and Otto Ruehr, in their solutions, rewrite the identity (1) in hypergeometric notation:

$$\frac{p}{p+1}\,{}_3F_2\left(\begin{matrix}1+p, & 1,1\\ 2+p, & 2\end{matrix}\,\middle|\,1\right)=p\,{}_3F_2\left(\begin{matrix}1-p, & 1,1\\ 2, & 2\end{matrix}\,\middle|\,1\right)$$

which is a special case of a more general form found on p. 14 of W. N. Bailey, *Generalized Hypergeometric Series*, Stechert–Hafner.

(2) A simpler form of the problem, proposed by Robert E. Shafer, was solved by R. G. Buschman in *Mathematics Magazine*, Problem 355, 1958, p. 107.

Majorizing Properties of Coefficients of Tchebychef Polynomials

6084 [1976, 292]. *Proposed by Theodore J. Rivlin, IBM, Yorktown Heights, New York*

Let $T_n(x)=t_0+t_1x+\cdots+t_nx^n$ denote the Tchebychef polynomial of degree n (that is, $T_n(x)=\cos n\theta$ where $x=\cos\theta$). Suppose that $p(x)=a_0+a_1x+\cdots+a_nx^n$ is real-valued. Show that if $|p(\cos(j\pi/n))|\leqq 1$ for $j=0,1,\ldots,n$ then

$$|a_{n-2m}|+|a_{n-2m-1}|\leqq|t_{n-2m}|,$$

$m=0,1,\ldots,[(n-1)/2]$. (The fact that $|a_{n-2m}|\leqq|t_{n-2m}|$ is well known.)

Solution by David G. Cantor, University of California, Los Angeles. We shall prove the following more general result:

THEOREM. *Let $\alpha_0>\alpha_1>\cdots>\alpha_n$ be $n+1$ distinct real numbers in $[-1,1]$ satisfying $\alpha_i+\alpha_{n-i}=0$ for $0\leqq i\leqq n$. Suppose $T_n(x)=\sum_{i=0}^n t_ix^i$ is a polynomial of degree n satisfying $T_n(\alpha_i)=(-1)^i$ for $0\leqq i\leqq n$. If $A(x)=\sum_{i=0}^n a_ix^i$ has real coefficients and satisfies $|A(\alpha_i)|\leqq 1$ for $0\leqq i\leqq n$, then $|a_{n-2m}|+|a_{n-2m-1}|\leqq|t_{n-2m}|$ for $0\leqq m<n/2$. (The problem is the case $\alpha_i=\cos(i\pi/n)$ for $0\leqq i\leqq n$.)*

Proof. Suppose $A(\alpha_i)=\gamma_i$, $0\leqq i\leqq n$. We shall obtain a formula expressing the a_i in terms of the γ_i. Put $u(x)=\prod_{i=0}^n(x-\alpha_i)$ and $c_i(x)=u(x)/((x-\alpha_i)u'(\alpha_i))$. Then, using standard results from interpolation theory, $A(x)=\sum_{i=0}^n\gamma_ic_i(x)$. Next, if $n=2r-1$, then $u(x)=\prod_{h=0}^{r-1}(x^2-\alpha_h^2)$, and hence

$$c_i(x)=\prod_{h=0}^{r-1}(x^2-\alpha_h^2)/((x-\alpha_i)u'(\alpha_i))$$

$$=(x+\alpha_i)\prod_{\substack{0\leqq h\leqq r-1\\ h\neq i,n-i}}(x^2-\alpha_h^2)/u'(\alpha_i);$$

if $n=2r$, then similarly

$$c_i(x)=\begin{cases} x(x+\alpha_i)\displaystyle\prod_{\substack{0\leqq h\leqq r-1\\ h\neq i,n-i}}(x^2-\alpha_h^2)/u'(\alpha_i), & \text{if } i\neq r\\[2ex] \displaystyle\prod_{0\leqq h\leqq r-1}(x^2-\alpha_h^2)/u'(\alpha_r), & \text{if } i=r.\end{cases}$$

Write $c_i(x)=\sum_{j=0}^n c_{i,j}x^j$. Now the products $\prod(x^2-\alpha_h^2)$ are monic polynomials in x^2, whose coefficients have alternating signs. Since $\operatorname{sign}(u'(\alpha_i))=(-1)^i$, it is easy to verify that $\operatorname{sign}(c_{i,n-2m})=(-1)^{i+m}$. Clearly $c_{i,n-2m-1}=\alpha_ic_{n-2m}$ and $c_{i,n-2m}=c_{n-i,n-2m}(-1)^n$. Then, if $n=2r-1$,

$$a_{n-2m}=\sum_{i=0}^n\gamma_ic_{i,n-2m}=\sum_{i=0}^{r-1}(\gamma_i-\gamma_{n-i})c_{i,n-2m},$$

and $a_{n-2m-1}=\sum_{i=0}^{r-1}\alpha_i(\gamma_i+\gamma_{n-i})c_{i,n-2m}$. Similarly if $n=2r$

$$a_{n-2m} = \sum_{i=0}^{r-1} (\gamma_i + \gamma_{n-i}) c_{i,n-2m} + \gamma_r c_{r,n-2m},$$

$$a_{n-2m-1} = \sum_{i=0}^{r-1} \alpha_i (\gamma_i - \gamma_{n-i}) c_{i,n-2m}.$$

Then, if $n = 2r - 1$

$$|a_{n-2m}| + |a_{n-2m-1}| \leqq \sum_{i=0}^{r-1} |c_{i,n-2m}| (|\gamma_i - \gamma_{n-i}| + \alpha_i |\gamma_i + \gamma_{n-i}|)$$

$$\leqq 2 \sum_{i=0}^{r-1} |c_{i,n-2m}| = \sum_{i=0}^{n} |c_{i,n-2m}|,$$

since all $|\gamma_i| \leqq 1$. The inequality $|a_{n-2m}| + |a_{n-2m+1}| \leqq \sum_{i=0}^{n} |c_{i,n-2m}|$ is obtained similarly, when $n = 2r$. To calculate the coefficients of $T_n(x)$ we substitute $\gamma_i = (-1)^i$ into the above formulas and obtain $|t_{n-2m}| = \sum_{i=0}^{n} |c_{n-2m}|$ and $t_{n-2m+1} = 0$. Thus

$$|a_{n-2m}| + |a_{n-2m+1}| \leqq |t_{n-2m}|.$$

Also solved by the MIT Combinatorics Class, and by the proposer.

Majorants for Families of Uniformly Integrable Functions

6085 [1976, 292]. *Proposed by William J. Sánchez, Courant Institute of New York University*

Call a family F of functions uniformly integrable (UI) if there exists $k(\varepsilon)$ such that $\int\{|f|d\mu : |f| > k\} < \varepsilon$ for all $f \in F$. If there exists integrable h such that $|f| \leqq h$ (a.e.) for all $f \in F$, then F is UI (Burrill, *Measure, Integration, and Probability*, p. 183). Is the converse true?

I. *Solution by Louis H. Blake, Richmond College of CUNY.* The converse to the stated theorem is not true even if we restrict ourselves to a discrete probability space.

Let $\Omega =$ the set of positive integers and $P(n) = 1/2^n$. Let

$$f_n(x) = \begin{cases} 2^n/n & \text{if } x = n, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\{f_n\}_{n \geqq 1}$ is a uniformly integrable family. The smallest measurable function h dominating $\{f_n\}_{n \geqq 1}$ is $h(n) = 2^n/n$ for each n and h is not integrable.

II. *Solution by Bertram Walsh, Rutgers University.* Let $f \in L^1 \cap L^p$, define q to be $1/p + 1/q = 1$, and let $S(f, k) = \{x : |f(x)| > k\}$. Then the inequality

$$\|f\|_1 \geqq \int_{S(f,k)} |f| d\mu \geqq k \cdot \mu(S(f, k))$$

yields $\|\chi_{S(f,k)}\|_q = [\mu(S(f, k))]^{1/q} \leqq \|f\|_1^{1/q}/k^{1/q}$. By Hölder's inequality

$$\int_{S(f,k)} |f| d\mu = \int |f| \chi_{S(f,k)} d\mu \leqq \|f\|_p \|\chi_{S(f,k)}\|_q \leqq \frac{\|f\|_1^{1/q} \|f\|_p}{k^{1/q}}.$$

It follows that if $1 < p \leqq \infty$, then any set $\{f \in L^1 \cap L^p ; \|f\|_1 \leqq K, \|f\|_p \leqq M\}$ is uniformly integrable, e.g., the unit ball in $L^p[0, 1]$ is uniformly integrable, but it is verified, by using multiples of characteristic functions of measurable sets, that for $1 < p < \infty$ no such unit ball is contained in a set $\{f \in L^1 : |f| \leqq h\}$ for any $h \in L^1$.

There is another way to look at the question. There is a theorem by J. L. Doob that the uniformly

integrable subsets of $L^1(\mu)$ are precisely those that are relatively compact in the weak topology of that Banach space. "Order intervals" like $\{f \in L^1: |f| \leqq h\}$ are weakly compact (essentially because they look like the unit ball of an L^∞-space in its weak*-topology); if μ is a totally finite measure so there is a natural mapping of $L^p \to L^1$ for $p > 1$, then the reflexivity of the L^p's makes their unit balls weakly compact in L^p and hence in L^1, therefore uniformly integrable. It can be shown that every uniformly integrable set is contained in the unit ball of some L^Φ, where Φ is an appropriate convex function with $\Phi(t)/t \to +\infty$ as $t \to +\infty$, so the example is "nearly exhaustive."

Also solved by Louis Blake, Claude Burrill, Roy Davies (England), Michael Ecker, Richard Enison, Le Baron Ferguson, Jerome Goldstein, Ellen Hertz, Barthel Huff (Canada), A. A. Jagers (Netherlands), Ole Jørsboe (Denmark), Joel Levy, Stuart Lloyd, O. P. Lossers, Jr. (Netherlands), John McCleary, MIT Combinatorics Class, Six problem solvers at the University of Santa Clara, Gideon Schectman (Israel), Randolph Schiefer, and Terence Shore.

REVIEWS

EDITED BY J. ARTHUR SEEBACH, JR. AND LYNN A. STEEN
with the assistance of the mathematics departments of St. Olaf and Carleton Colleges
COLLABORATING EDITOR FOR FILMS: SEYMOUR SCHUSTER, CARLETON COLLEGE

We invite readers to submit reviews of significant recent college-level mathematics books. We especially encourage reviews based on classroom use, or comparative reviews of several related books. Reviews should ordinarily not exceed two pages (per book) typed double spaced. Manuscripts of reviews as well as books submitted for review should be sent to: Book Review Editor, American Mathematical Monthly, St. Olaf College, Northfield, MN 55057.

Discrete Mathematical Models with Applications to Social, Biological, and Environmental Problems. By Fred S. Roberts. Prentice-Hall, Englewood Cliffs, New Jersey, 1976. xvi + 559 pp. $17.95. (Telegraphic Review, August-September 1976.)

Professor Fred Roberts presents us with a refreshingly new book full of a wide variety of well-developed examples. His book brings together a collection of topics in applied mathematics which had not previously been available in a textbook.

These topics all have a common thread, a connection with graphs and digraphs. However, the analysis presented includes many appropriate techniques and this book should not be considered only as a text of applied graph theory (although it would serve well in that capacity). The general emphasis of the book is reflected in the organization in chapters each of which deals with a particular mathematical technique illustrated with numerous applications. It is not simply a collection of different individual models as can be found, for example, in the short book by John Kemeny and J. Laurie Snell, *Mathematical Models in the Social Sciences*, but rather it makes a thorough development of applied techniques.

With this organization, Roberts' book is well suited for a course on applied mathematics. I recently used it, along with the above book by Kemeny and Snell, for a course on mathematics applied to the social sciences. Roberts' organization makes the chapters of his book highly independent, so that individual chapters, such as chapter 5 on Markov chains or chapter 7 on group decision-making, may be covered on their own. Thus it works very effectively with a complementary book with emphasis on specific models.

However, *Discrete Mathematical Models* stands on its own as an excellent text for applied mathematics at the advanced undergraduate or graduate level. A semester course might include the first introductory chapter, chapter 2 on graph theory, chapter 3 on the application of graphs, and any

combination of two or three other chapters. The fundamental ideas of graph theory presented in chapter 2 are necessary for chapters 3 and 4 and are helpful in the other chapters. The other chapters are largely self-contained so that very different courses can be taught from this text. The book would also serve well as a text for a full year course. Most of the material in all the chapters could be covered without any slack time.

Roberts' style is quite lucid and readable. He develops the mathematics thoroughly, but often leaves longer or deeper proofs for a separate section. This facilitates the use of the book for a course emphasizing applications without sacrificing the precision needed for a more thorough development.

The thought and care that has gone into Roberts' writing can be illustrated by some examples. In chapter 3, he explicitly reinforces the ideas on the "model-building cycle" presented in chapter 1 in a very natural way. In chapter 4, during the development of theorems on the eigenvalues of adjacency matrices of digraphs and the properties of pulse processes, a very nice application of linear algebra, Roberts very smoothly emphasizes the use of the contrapositive form of a theorem, a point which can often be confusing to students. Chapter 7 includes the most readable, reasonably rigorous treatment of Arrow's impossibility theorem which I have seen. Some students in my applied mathematics class developed projects by independently reading sections of the book not covered in class and working the associated problems.

The mixture of applications and theory is uneven throughout the book. Chapters 2, 3, and 4 present a neat unit, with the basic graph theory in 2, applications in 3 and 4. The validity of the models using pulse processes presented in chapter 4 have been a subject of some discussion in this MONTHLY (Roberts and Brown, 82 (1975) pp. 577–593, and Waterhouse, 84 (1977) pp. 25–27), but seem to be adequately couched with qualifications in this book. Chapter 5 itself is a similar unit, with the first six sections on the theory and the last five on specific applications. In this chapter, in particular, as well as in section 7.4, Roberts' acknowledged debt to Kemeny and Snell and their aforementioned book is evident. The last three chapters are distinctly different. Chapters 6 and 8 on n-person game theory and utility and measurement especially emphasize the development of the theory, with examples only lightly treated and further developed in the exercises. There is little discussion of the implications the theory might have as, for example, in the application of game theory to oligopoly.

A text such as this must have adequate exercises. Roberts has done a superior job in providing a large number of exercises which range from routine exercises on techniques to proofs omitted from the text to extensions of models in new directions. The latter sorts of problems can serve as the basis for a term project. A student who works through a selection of these exercises must develop some understanding of the material.

Discrete Mathematical Models should serve as a standard for texts on applications of discrete mathematics. It is suitable for use in a variety of courses at the advanced undergraduate and graduate levels and will, indeed, probably lead to the development of new courses in this field. Roberts has made an excellent contribution to the teaching of applied mathematics.

CHRISTOPHER H. NEVISON, Colgate University

TELEGRAPHIC REVIEWS

Telegraphic reviews are designed to give prompt notice of new books with sufficient information to assist our readers in deciding whether to order an examination copy or to suggest library purchase. Possible uses are indicated as follows:

T = textbook | P = professional reading
S = supplementary reading | L = undergraduate library purchase
13 to 18 = freshman to second year graduate level usage
1 to 4 = appropriate time in semesters to cover text

Asterisks (*) or question marks (?) denote special positive or negative emphasis, respectively. Publishers are denoted by standard abbreviations; complete addresses may be found in *Books in Print*.

PRECALCULUS, T(13: 1). *Precalculus: A Functional Approach with Applications*. Salvatore Barbasso, John Impagliazzo. HarBrace J, 1977, xiv + 413 pp, $12.95. The material is built around the concept of a function. Chapters on linear, quadratic, circular, exponential, logarithmic, polynomial, and rational functions. Also includes complex numbers, induction, sequences, and series. Attractive text format. Appendices. Tables. Exercises and review exercises. Answers to selected exercises. Index.RJA

EDUCATION, T(13: 1, 2). *Modern Mathematics, An Elementary Approach, Fourth Edition*. Ruric E. Wheeler. Brooks/Cole, 1977, xx + 679 pp, $14.95. For elementary school teachers and liberal arts students, this edition de-emphasizes the structure of mathematics and proof-writing while adding applications of mathematics and a chapter on computers and programming. (*Third Edition*, TR, October 1973). JNC

EDUCATION, P. *Overview and Analysis of School Mathematics, Grades K-12*. NCTM, 1977, xiv + 157 pp, $3 (P). Reprinting by NCTM of the 1975 CBMS study (TR, January 1976) by the National Advisory Committee on Mathematics Education. LAS

HISTORY, P, L*. *The Correspondence of Isaac Newton, V. VI, 1713-1718*. Ed: A. Rupert Hall, Laura Tilling. Cambridge U Pr, 1976, xxxviii + 499 pp, $65. Letters and other brief notes from Newton's eighth decade. Many pertain to business of the Mint, but most of the source documents of the Newton-Leibniz controversy are here too. The introduction includes a detailed reference chart outlining the chronology of this dispute. LAS

HISTORY, P, L**. *Was Pythagoras Chinese? An Examination of Right Triangle Theory in Ancient China*. Frank J. Swetz, T.I. Kao. NCTM, 1977, 75 pp, $4.40 (P). An extensively annotated translation of the ninth chapter of the *Chin chang suan shu*, the most important mathematical text from ancient China (probably early Han dynasty, third century B.C.). The ninth chapter, titled *Kou ku* (right angles) presents 24 problems concerning properties of right triangles, including many applications of the "Pythagorean" theorem. LAS

COMBINATORICS, P. *Lecture Notes in Mathematics-558: On Construction and Identification of Graphs*. Boris Weisfeiler. Springer-Verlag, 1976, xiv + 237 pp, $11 (P). Graph identification is the problem of determining whether two graphs are isomorphic. Exposition is highly algebraic, treating cellular algebras which have wide application in combinatorics and finite group theory. Emphasizes algorithms and use of computers. SS

NUMBER THEORY, P?. *Congruence Surds and Fermat's Last Theorem*. Max M. Munk. Vantage Pr, 1977, 33 pp, $7. A seventeen page proof of Fermat's Last Theorem which the author claims requires powerful mental resonance in order to understand. Clearly, the author's mind is in a state of resonance which the reviewer was not able to attain. CEC

ALGEBRA, T(16-17: 1), L. *First Course in Rings, Fields and Vector Spaces*. P.B. Bhattacharya, S.K. Jain. Halsted Pr, 1977, ix + 238 pp, $6.95. A sequel to the authors' *First Course in Group Theory* (TR, October 1973). Rings and modules through the Wedderburn-Artin, Goldie-Lesieur-Crosoit, and Krull-Schmidt theorems, and fields through the fundamental theorem of Galois theory. Basic linear algebra included. No bibliography. Prerequisite: groups and maturity. TRS

DIFFERENTIAL EQUATIONS, T(16-18: 1), S, P, L. *The Finite Element Method in Partial Differential Equations*. A.R. Mitchell, R. Wait. Wiley, 1977, x + 198 pp, $14.95. Text on the numerical solution of partial differential equations using the finite element method. Advanced calculus and vector spaces are the prerequisites. Chapters on variational principles, methods of and convergence of approximations, basis functions, time-dependent problems, and applications. Exercises. References. Index. RJA

ALGEBRAIC GEOMETRY, P, L. *Zeta-Functions: An Introduction to Algebraic Geometry*. A.D. Thomas. Pitman, 1977, vii + 230 pp, $15 (P). An introduction to the statements and proofs of the Weil conjectures for varieties over a finite field. While a great deal of mathematics is presumed and much is left unproved, these notes *motivate* and describe the successful attack on the Weil conjectures. The analogies between number fields and function fields, and between varieties over C and over finite fields is emphasized for motivational purposes. An extremely valuable book. SG

TOPOLOGY, T(18), P. *Introduction to Fibre Bundles*. Richard D. Porter. Pure and Appl. Math., V. 31. Dekker, 1977, v + 170 pp, $19.50 (P). This typewritten set of lecture notes gives a concise introduction to the theory of fibre bundles, emphasizing the classical groups. The approach is heavily algebraic. Includes cellular decomposition of O(n), U(n), Sp(n), the covering homotopy theorems, and construction of the universal bundles for the above groups. Good problems, index and bibliography.TLS

STATISTICS, T*(13-14: 1). *Statistics: An Introduction*. Harold J. Larson. Wiley, 1975, ix + 418 pp, $13.25. Noncalculus version. A quick entry into inferential statistics made possible by putting counting problems and elementary probability in the appendix. A unique feature of the text is a major emphasis on expected value and the difference between discrete and continuous random variables. Large number of exercises, chosen from social and physical sciences. LCL

APPLICATIONS (DEMOGRAPHY), S(15-17), P, L*. *Applied Mathematical Demography.* Nathan Keyfitz. Wiley, 1977, xxiv + 388 pp, $19.95. A clear, interesting attempt "to find answers that will be serviceable to those working on population and related matters." Makes extensive but straightforward use of elementary calculus and linear algebra, thereby providing a rich source of interesting applications for elementary mathematics courses. Contains many simple arguments leading to surprising conclusions, e.g., a decrease in the population growth rate of .01 results in an increase of about 30% in premium for a contributory nonreserve pension. A good resource for a seminar in mathematical modelling. LAS

APPLICATIONS (ENERGY), P. *Mathematical Aspects of Production and Distribution of Energy.* Peter D. Lax. Proc. of Symp. in Appl. Math., V. 21. AMS, 1977, v + 137 pp, $14.40. Seven papers from the Energy Short Course held at the San Antonio meeting of the AMS in January, 1976, in two groups: models of energy production employing finite difference and finite element methods on systems of partial differential equations, and models of energy distribution employing statistics, optimization and systems analysis. LAS

APPLICATIONS (FLUID DYNAMICS), P. *Proceedings of the International Symposium on Modern Developments in Fluid Dynamics.* Ed: J. Rom. SIAM, 1977, xviii + 393 pp, $37.50. Papers in honor of Sydney Goldstein, including a survey of Goldstein's role in fluid dynamics by Sir James Lighthill. LAS

APPLICATIONS (MEDICINE), P, L. *Environmental Health, Quantitative Methods.* Ed: Alice Whittemore. SIAM, 1977, vii + 259 pp, $19.50 (P). 17 papers from a July 1976 SIMS (Siam Institute for Mathematics in Society) conference at Alta, Utah, in which mathematicians, statisticians and epidemiologists gathered to discuss quantitative analysis of environmentally induced disease. LAS

APPLICATIONS (ENGINEERING), P. *Lecture Notes in Mathematics-565: Turbulence and Navier Stokes Equations.* Ed: Roger Temam. Springer-Verlag, 1976, 194 pp, $8 (P). The lectures from the workshop of June 1975 held at University de Paris-Sud à Orsay, France. JAS

APPLICATIONS (PHYSICS), S(16-18). *Problem Book in Relativity and Gravitation.* Alan P. Lightman, *et al.* Princeton U Pr, 1975, xiv + 599 pp, $20. A collection of almost 500 problems with solutions and explanations dealing with special relativity, general relativity, and cosmology. JAS

APPLICATIONS (PHYSICS), P*, L. *An Exposition of Catastrophe Theory and Its Applications to Phase Transitions.* D. O'Shea. Pure and Appl. Math., No. 47. Queen's U, ix + 200 pp, (P). An economical classification of phase transition diagrams for pure and mixed liquids, derived by applying elementary catastrophe theory to a single, simple premise relating the mathematics to the physics. Begins with a careful, extensive survey of catastrophe theory, including an intuitive treatment of Thom's classification theorem. LAS

APPLICATIONS (PHYSICS), P. *Foundations of the Mathematical Theory of Structures.* E.R. de Arantes e Oliveira. Springer-Verlag, 1975, 223 pp, $15.20 (P). Convergence theorems that extend the variational methods used to characterize solutions of differential equations to cover the finite element method as well, followed by applications of these methods to the theory of structures. LAS

APPLICATIONS (PHYSICS), T(16: 1), S, L. *Vectors and Tensors for Engineers and Scientists.* Fred A. Hinchey. Halsted Pr, 1976, xi + 298 pp, $9.95. A solid, traditional course in vector analysis which requires a background in physics. There are lots of problems and helpful diagrams. CEC

APPLICATIONS (PHYSICS), T(15-16: 2). *Vector Fields.* J.A. Shercliff. Cambridge U Pr, 1977, xi + 329 pp, $28.50; $7.95 (P). "An expanded version of thirty second-year lectures--for students of engineering" with a reasonable number of problems and a good index. A interesting looking presentation of classical physics. JAS

APPLICATIONS (PHYSICS), T(17: 1), S, P, L. *Mechanics of Swimming and Flying.* Stephen Childress. Courant Inst, 1977, v + 156 pp, $4.75 (P). The fluid dynamics encountered in the study of the swimming of microorganisms and fish, and the flying of birds and insects, are studied via Navier-Stokes equations. Exercises and an extensive bibliography are included. CEC

APPLICATIONS (PHYSICS), P. *Lecture Notes in Physics-64: Waves on Water of Variable Depth.* Ed: D.G. Provis, R. Radok. Springer-Verlag, 1977, 231 pp, $11.50 (P). 26 papers from a July 1976 international symposium held in Canberra, Australia. Papers are grouped in sections (e.g., tsunami generation, waves on beaches, long period waves), each introduced by an editor's summary. LAS

APPLICATIONS (SEISMOLOGY), T(17: 2), S. *Some Mathematical Topics in Seismology.* Robert Burridge. Courant Inst, 1976, viii + 317 pp, $9.75 (P). Topics include elasticity theory, plane and surface waves, discontinuities in the tangential component of displacement, Lamb's problem, ray theory, Backus-Gilbert theory, fracture mechanics and seismic sources. Assumes only a good undergraduate applied mathematics background. Drawings, tables, diagrams, many references. Reproduced from typescript. DFA

APPLICATIONS (SOCIAL SCIENCE), P?. *Combinatorial Connectivities in Social Systems.* R.H. Atkin. Birkhäuser, 1977, ii + 239 pp, sFr. 38 (P). Exposition of the author's eccentric endeavors to apply exterior algebras and varieties of homotopy to social organization within the University of Essex. Purports to suggest new possibilities of determinism in understanding the behavior of systems. Cites only the author's related studies for documentation. LAS

Reviewers Whose Initials Appear Above

Richard J. Allen, St. Olaf; David F. Appleyard, Carleton; Judith N. Cederberg, St. Olaf; Clifton E. Corzatt, St. Olaf; Steven Galovich, Carleton; Loren C. Larson, St. Olaf; Thomas R. Savage, St. Olaf; Seymour Schuster, Carleton; J. Arthur Seebach, Jr., St. Olaf; Lynn Arthur Steen, St. Olaf; Timothy L. Strotman, St. Olaf.

NEWS AND NOTICES

EDITED BY RAOUL HAILPERN, SUNY at Buffalo

Readers are invited to contribute to the general interest of this department by sending news items to Mathematical Association of America, 1225 Connecticut Avenue, NW, Washington, D.C. 20036. Items must be submitted at least five months before publication can take place.

PERSONAL ITEMS

University of North Carolina, Chapel Hill: Mr. T. T. Trent, University of Virginia, has been appointed Instructor; Associate Professors R. B. Gardner and L. D. Geissinger have been promoted to Professors.

University of North Carolina, Wilmington: Associate Professors Thaddeus Dankel and Fletcher Norris have been promoted to Professors.

Wellesley College: Assistant Professors Ann K. Stehney and Alan Shuchat have been promoted to Associate Professors.

Assistant Professor John Grant, University of Florida, has been promoted to Associate Professor.

Professor William Floyd Hill, East Texas State University, retired on July 31, 1977, with the title of Emeritus Professor.

Associate Professor Eric Langford, University of Maine at Orono, has been promoted to Professor.

Professor D. A. Sanchez, UCLA, has been appointed Professor at the University of New Mexico.

E. M. Scheuer, Professor of Management Science and Professor of Mathematics at California State University, Northridge, will spend the 1977–78 academic year on sabbatical leave as Visiting Professor of Statistics in the Department of Mathematics at the City University, London.

Assistant Professor Kathleen A. Taylor, Duquesne University, has been promoted to Associate Professor.

Professor Olga Taussky Todd, California Institute of Technology, retired on July 1, 1977, with the title of Professor Emeritus.

Brother Damian Connelly, Professor at La Salle College, died on April 25, 1977, at the age of 58. He was a member of the Association for four years.

Professor Winston L. Massey, University of Tennessee at Chattanooga, died on February 14, 1977, at the age of 72. He was a member of the Association for forty years.

Ms. Mary McKenna, New York City, died in November 1976. She was a member of the Association for forty-one years.

EXAMINATIONS IN MATHEMATICS FOR SECONDARY SCHOOLS

The Annual High School Mathematics Examinations Prior Year Examinations 1969–1977. Specimen sets of prior year examinations containing a question booklet, a solution key and a copy of the pamphlet "How About A Career With Mathematics?" are available for 35¢ each.

The USA and International Mathematical Olympiads 1976–1977. These pamphlets contain the problems and solutions to the 1976 and 1977 Olympiads. Price: 1976—35¢ each; 1977—50¢ each.

Summary of Results and Awards 1969–1977. This publication describes the results and awards from prior Annual High School Mathematics Examinations. The 1977 edition is based on reports received from some 6000 participating schools and 340,000 registered students. All prior year Summaries are available from 1969–1977 for $1 each.

These publications may be obtained from: Dr. W. E. Mientka, Executive Director, Annual High School Mathematics Examination, Department of Mathematics and Statistics, 917 Oldfather Hall, University of Nebraska, Lincoln, NE 68588.

A minimum order of $1.20 is required. Advance payment must accompany each order. All prices include postage and handling. Make check payable to MAA COMMITTEE ON HIGH SCHOOL CONTESTS.

MATHEMATICAL ASSOCIATION OF AMERICA

Official Reports and Communications

BY-LAWS OF THE MATHEMATICAL ASSOCIATION OF AMERICA (INC.)

ARTICLE I — NAME, PURPOSE AND CORPORATE SEAL

1. This organization shall be known as

THE MATHEMATICAL ASSOCIATION OF AMERICA (INCORPORATED)

2. Its object shall be to assist in promoting the interests of the mathematical sciences in America, especially in the collegiate field, by holding meetings in any part of the United States or Canada for the presentation and discussion of mathematical papers, by the publication of mathematical papers, journals, books, monographs, and reports, by conducting investigations for the purpose of improving the teaching of mathematics, by accumulating a mathematical library and by cooperating with other organizations whenever this may be desirable for attaining these or similar objects.

3. The Corporate Seal of the Association shall have inscribed thereon the name of the Association and the words "Corporate Seal — Illinois".

ARTICLE II — MEMBERSHIP

1. There shall be two classes of members: individual and institutional.

2. Any person interested in the field of collegiate mathematics shall be eligible for election to individual membership in the Association.

3. Any institution, academic or corporate, interested in the support of collegiate mathematics shall be eligible for election to institutional membership in the Association.

4. Election to membership shall be by vote of the Board upon written application from the individual or institution seeking admission. In the case of individuals qualifying for student dues, the application shall be endorsed by two individual members of the Association.

ARTICLE III — BOARD OF GOVERNORS AND OFFICERS

1. The Officers of the Association shall be a President, a President-Elect (only during a year immediately prior to the expiration of a President's term), a Past-President (only during a year immediately following the expiration of a President's term), a First Vice-President, a Second Vice-President, an Editor of its publication entitled "THE AMERICAN MATHEMATICAL MONTHLY", a Secretary, and a Treasurer.

2. There shall be a Board of Governors (herein called "the Board") to consist of the officers, the ex-presidents for terms of six years after the expiration of their respective presidential terms, the Editor of each of its two publications entitled TWO-YEAR COLLEGE MATHEMATICS JOURNAL and MATHEMATICS MAGAZINE, the members of the Finance Committee, and additional elected members (herein called "Governors"). It shall be the function of the Board to supervise all scholarly and scientific activities of the Association, to administer and control these activities, and to authorize expenditures of funds of the Association.

3. There shall be an Executive Committee of the Board consisting of the President, the President-Elect (only during a year immediately preceding the expiration of a President's term), the Past-President (only during a year immediately following the expiration of a President's term), the two Vice-Presidents, the Editor of the AMERICAN MATHEMATICAL MONTHLY, the Secretary, and the Treasurer. It shall be the function of this Committee to review continually the policies and activities of the Association, to plan and organize new activities, to formulate in broad outline the programs of meetings and of publications, and in general to consider all matters of importance or interest to the Association. This Committee shall prepare the agenda for meetings of the Board and shall analyze the implications and aspects of all matters which are to come before the Board for decision. It shall present to the Board the viewpoints suggested by such analyses, as well as all such facts as may seem pertinent or as may in any way facilitate the Board's work.

4. At all meetings of the Board of Governors a quorum shall consist of not less than 25 per cent of the membership of the Board, and no business may be validly transacted at a meeting at which less than a quorum is present.

5. There shall be a Finance Committee responsible to the Board; at the direction of the Board it shall receive and administer the funds of the Association, control its properties and investments, make its contracts, and exercise such powers as may be delegated to it by the Board. This Committee shall consist of six members including the President, the Past-President (for a term of two years), the Secretary, and the Treasurer.

6. The Board shall hold a meeting each year immediately preceding the annual business meeting of the Association. Other meetings of the Board may be held from time to time at the call of the President or of any six (6) members of the Board.

7. Notice of all meetings of the Board shall be given by the Secretary to each member of the Board at least fifteen (15) days prior to the date set therefor.

8. A member of the Board may waive notice with the same effect as if due notice had been given.

9. The Board may refer a matter to a referendum mail vote of the entire membership and shall make such reference if a referendum is requested, prior to final action by the Board, by three hundred or more members. The taking of a referendum shall act as a stay upon Board action until the votes have been canvassed, and thereafter no action may be taken by the Board except in accordance with a plurality of the votes cast in the referendum.

ARTICLE IV — ELECTIONS, APPOINTMENTS, AND TERMS OF OFFICERS AND MEMBERS OF THE BOARD

1. (a) The membership at large shall elect biennially a President-Elect for a term of one year and a First Vice-President for a term of two years. The President-Elect shall become President for a two-year term at the expiration of the one-year term as President-Elect and shall become Past-President for a one-year term at the expiration of the term as President.

(b) The membership in each Section shall elect triennially a Governor for a term of three years beginning July 1. For these elections at least two nominations shall be submitted to the members by a committee appointed for that purpose by the Chairman of the Section. A Governor who has moved his or her place of employment from the Section by which he or she was elected shall be considered to have ended his or her term of office on the Board.

(c) The Board shall elect annually two Governors for terms of three years and at appropriate times by ballot and for terms stated: a Second Vice-President for two years, an Editor of the AMERICAN MATHEMATICAL MONTHLY, an Editor of the TWO-YEAR COLLEGE MATHEMATICS JOURNAL, an Editor of MATHEMATICS MAGAZINE, a Secretary, and a Treasurer each for five years, and members of the Finance Committee (other than the President, the Past-President, the Secretary, and the Treasurer) for four years.

(d) The beginning and end of the term of every officer and member of the Board (except as provided in Section (b) of this Article) shall occur at the adjournment of the annual business meeting. All officers and members of the Board shall hold over until their respective successors have been duly elected or appointed and qualified.

(e) The President shall be ineligible for reelection as President-Elect or as President. The Vice-Presidents, the Editors, and the Governors shall be eligible for reelection only after an interim equal to their respective terms of office except that Governors having served less than a year and a half shall be eligible for reelection for a term of three years.

(f) The Board shall have authority to fill vacancies *ad interim* in any office, including vacancies in the Board, and to make any other appointments necessary for the transaction of business of the Association.

(g) Elections by the Board shall be made from nominations by the Executive Committee. At least two nominations shall be made for each office to be filled in the case of the Second Vice-President, Governors (except Sectional Governors), and members of the Finance Committee. The Board may make additional nominations.

2. For general elections by the membership of the Association there shall be a Nominating Committee appointed annually by the President with the approval of the Board. The general election shall be conducted in two stages, a primary mail voting concluding approximately five months before the date of the annual meeting and a final voting concluding at the time of the annual meeting. For the primary voting the Nominating Committee shall prepare printed ballots with five or more nominees for President-Elect and three or more for each other office to be filled by the members. Blank spaces on the ballot shall be provided for write-in votes. From the results of the primary voting the Nominating Commitee shall prepare a printed ballot for the final voting. This ballot shall be mailed to the membership approximately one month before the annual meeting and the voting shall close on the day of the annual business meeting. The final ballot shall present one nominee for President-Elect, to be selected by the Nominating Committee out of the three persons who received the most votes in the primary voting. For

each other office the final ballot shall present two names, one being that of the person who received the highest vote in the primary voting.

3. The President shall be the Executive Officer of the Association, shall preside at all meetings of the Board of Governors and at the annual business meeting of the Association, shall be Chairman of the Executive Committee and of the Finance Committee; and shall have the usual duties pertaining to the office and such other duties as may from time to time be assigned by the Board of Governors.

4. In the absence of the President, the First Vice-President (or in his or her absence the Second Vice-President) shall have and exercise the powers of the President, except that the Past-President shall preside at meetings of the Finance Committee (or in his or her absence the senior member, in terms of service on the Committee, of the two elected members of the Finance Committee). The Board of Governors may assign to the Vice-Presidents such duties as may from time to time be determined.

5. The Secretary shall have the usual duties pertaining to the office, including the custody of the records of the Association and of its Corporate Seal, the keeping of minutes of the meetings of the Board of Governors and of the annual business meeting and special meetings, and the giving of due notice of all regular and special meetings of the Association and of the Board of Governors. The Secretary shall also have the duty of seeing that whenever Governors are elected, including the election of Governors to fill vacancies, a Certificate, under the Seal of the Association, giving the names of those elected and the terms of their office, shall be recorded in the Office of the Recorder of Deeds for Cook County, Illinois. Such Certificates shall be signed by the Secretary and verified by oath of the President.

6. The Treasurer shall have the usual duties pertaining to the office including the collection of dues and the supervision and safekeeping of the funds of the Association.

7. (a) There shall be an Executive Director who shall be a paid employee of the Association. He shall have charge of the central office of the Association and shall carry out such other duties as may be assigned to him by the Board. He shall be responsible to the Board and shall attend meetings of the Board, the Executive Committee, and the Finance Committee, except when they meet in executive session, but he shall not be *ex officio* a member of these bodies. He shall be especially responsible for implementing and coordinating Section activities.

(b) The Executive Director shall be elected by the Board under terms and conditions of employment fixed by the Finance Committee.

Article V — Business Meetings of the Association

1. A business meeting of the Association shall be held annually, at such time and place as the Board may direct. Other business meetings of the Association may be called from time to time by the Board or by the President of the Association to be held at such time and place as may appear from the call.

2. Notice of any business meeting of the Association shall be given by the Secretary to each member of the Association at least thirty (30) days prior to the date set for each meeting.

3. Any member of the Association may waive notice with the same effect as if due notice had been given.

4. At all business meetings of the Association a quorum shall consist of not less than twenty-five (25) members and no business may be validly transacted at a meeting at which less than a quorum is present.

Article VI — Sections

1. In the interest of more effective promotion of the objectives of the Association on a local level, the United States, Canada and their possessions shall be subdivided by the Board of Governors into non-overlapping geographical areas, and a Section of the Association shall be established in each of these areas. The subdivision into non-overlapping areas may be changed by the Board, upon recommendation by the Committee on Sections (see paragraph 7).

2. Each member of the Association residing in the United States, Canada or their possessions shall belong to one and only one Section. He will belong to the Section in whose geographic area he resides, except that a member who resides in one area and is employed in a different area may elect the Section to which he prefers to belong. Any member may petition the Committee on Sections for reassignment of his membership to another Section.

3. Each Section shall adopt a set of By-Laws which, along with any subsequent changes, must be approved by the Board. The geographic area covered by a Section shall be described in the By-Laws for the Section.

4. If there are members of the Association residing in a geographic area in which no Section has been

organized, any ten or more members of this Association residing or employed in this area may petition the Board for authority to organize a Section covering that area.

5. A group of not less than twenty-five members of an existing Section may petition the Board to partition the area and the Section into two or more Sections. The Board shall have authority to approve or deny this petition. The Board may specify conditions under which such action may be accomplished. It may conduct a poll of some or all members of the Association in the Section to determine the advisability of allowing the proposed partition of the Section. If separate Sections are approved then each new Section must prepare its own set of By-Laws to be approved by the Board.

6. A group of not less than twenty-five members residing or employed in that part of the area of an existing Section which they desire to become part of another existing Section may petition the Board to redefine the geographic boundaries of the Sections affected. The Board shall have authority to approve or deny this petition. It may conduct a poll of some of all members of the Sections involved to determine the advisability of permitting such action.

7. There shall be a standing Committee on Sections through which the Board shall maintain general supervision over the activities of all Sections. This Committee, in particular, shall study all matters involving creation of Sections or modification of boundaries of Sections and make appropriate recommendations to the Board.

8. The Association shall not be obligated to pay from its treasury any of the expenses of a Section except as the Board may provide.

Article VII — Official Publications

1. The Association shall publish at least one official journal, of which one shall be sent free to all members of the Association in accordance with Article VIII.

2. The Board shall have full control of the publication and sale of each official journal and of all other official publications.

3. There shall be appointed by the Board a body of Associate Editors for each official journal.

4. The Board shall from time to time, as the need arises, make special provision for the management of any other publications.

5. The Board shall fix the price of each official journal and of any other publications of the Association.

Article VIII — Dues

1. The Board shall establish the annual dues and privileges of membership for individual and institutional members. The dues of individual members shall include a subscription to one of the official journals.

2. All dues shall be payable on the first of January of each year. Should the annual dues of any member remain unpaid beyond a reasonable time, that member shall be dropped from the list after due notice.

3. New members entering the Association after April 1 of any year may have their dues prorated for the balance of the year, except when they desire to receive the full current volume of an official journal.

4. Any individual member who because of age is no longer in active service, who is in good standing at the time of retirement, and who has been a member of the Association for twenty years, may, upon notifying the central office of said retirement, be exempt from the payment of dues, with the privilege of obtaining an official journal at an annual cost of half of the dues of an individual member.

Article IX — Amendments to the Articles of Association and By-Laws

1. Changes in the Articles of Association or amendments to the By-Laws may be made at any annual business meeting of the Association, or at any adjourned session thereof, or at any special meeting of the Association called for such purpose, by a two-third (2/3) vote of those present and entitled to vote; *provided* that due notice concerning such amendment shall have been printed in each official journal, or mailed to each member, at least one (1) month before the date of such meeting. The Secretary shall give such due notice when so instructed by a vote of the Board of Governors or when so petitioned by at least forty members of the Association.

2. No changes in the Articles of Association or amendments to these By-Laws shall have legal effect until a certificate thereof, verified by oath of the President and under Seal of the Association, attested by the Secretary, shall be recorded in the office of the Recorder of Deeds for Cook County, Illinois.

MATHEMATICAL ASSOCIATION OF AMERICA
THE FIFTY-SEVENTH SUMMER MEETING OF THE ASSOCIATION

The Fifty-Seventh Summer Meeting of the Association was held at the University of Washington, Seattle, Washington, from Sunday, August 14, to Tuesday, August, 16, 1977, in conjunction with meetings of the American Mathematical Society, the Association for Women in Mathematics, the Conference Board of the Mathematical Sciences, the Institute of Mathematical Statistics, the Mathematicians Action Group, and Pi Mu Epsilon. The Pacific Northwest Section of the MAA also met in conjunction with the Association and held its Business Meeting at 4:30 P.M. on Tuesday. There were registered 1132 persons, including 665 members of the Association.

Sessions of the Association were held on Sunday morning and afternoon, Monday morning, and Tuesday afternoon. The Tuesday afternoon session was a dual session with one session emphasizing the subject matter of mathematics and the other emphasizing some classroom aspects. The latter session was organized by Professor Larry A. Curnutt of Bellevue Community College and was held in 130 Kane Hall. All other sessions were held in Meany Hall.

Presiding officers at the three Earle Raymond Hedrick Lectures by Professor Joseph B. Keller were President Henry L. Alder, First Vice-President R. Creighton Buck, and Second Vice-President Howard E. Zink; at the lecture by Dr. Brockway McMillan, Dr. Henry O. Pollak; at the lectures by Professors Jean J. Pedersen and Eileen L. Poiani, Professor Alice T. Schafer; at the lecture by Professor Isaac J. Schoenberg, Professor Edwin F. Beckenbach; at the lecture by Professor Dorothy Maharam Stone, Professor Kenneth A. Ross; at the lecture by Professor Bill Leonard, Professor Donald J. Albers; at the lecture by Professor David G. Larman, Professor Branko Grunbaum; at the lecture by Professor Donald W. Bushaw, Professor Calvin A. Lathan; at the lecture by Professor George E. Andrews, Professor Emma Lehmer; at the lectures by Professors Robert J. Bitts and Ben J. Jones, Professor Larry A. Curnutt.

FIRST SESSION OF THE ASSOCIATION

Welcome on behalf of the University of Washington by Professor Jack Segal, Chairman, Department of Mathematics, University of Washington.

The Earle Raymond Hedrick Lectures: *Mathematical Aspects of Athletics and of Vision*, Lecture I: *Athletics*, by Professor Joseph B. Keller, Courant Institute of Mathematical Sciences, New York University, on leave at Stanford University.

A variety of different athletic events were analyzed mathematically. These included weight lifting, tennis scheduling, the baseball pennant race, rowing, running, skiing, rope climbing, etc. In each case the relevant rules and mechanical or other principles were explained and in some cases the results of the analysis were compared with the records. Various kinds of mathematics were used, such as dimensional analysis, Latin squares and optimal control theory.

Panel Discussion: How to Teach Mathematics

A panel discussion with presentations by Professor Peter J. Hilton, Battelle Memorial Institute and Case Western Reserve University and Professor George Pólya, Stanford University. Moderated by Professor Jean J. Pedersen, University of Santa Clara.

What Not To Do (A Demonstration), by Professor Hilton.

Various standard errors committed in undergraduate mathematics lectures were exemplified vividly by Professor Hilton.

What To Do (Some Rules of Thumb), by Professor Pólya.

Two (rather simple, rough and ready) rules were offered and illustrated.

A rule for the *presentation* in class: <u>Stress the Familiar</u> (the uncomplicated, the concrete, the visible, the tangible).

A rule for the *preparation* of the lesson: <u>Try to See the Intuitive</u> (clear, distinct, easily visible, immediately convincing) aspect of the topic.

SECOND SESSION OF THE ASSOCIATION

Hedrick Lecture II: Color Vision, by Professor Joseph B. Keller.

Color matching and color mixing experiments and their results were described, as well as the

mathematical characterization of a colored light. Then the space of color sensations was introduced and its properties were determined to account for the experimental results. This leads to the color-brightness cone in a three-dimensional vector space. A metric is introduced into this space, and it is used to account for the color limen, or just noticeable difference between nearly identical sensations. The uniqueness of the metric was considered. In the course of the development, the concept of the transitive part of a binary relation was introduced to overcome the lack of transitivity of the relation "is distinguishable from".

The WAM Program: Women and Mathematics, by Professors Jean J. Pedersen, University of Santa Clara and Eileen L. Poiani, St. Peter's College.

Professor Poiani presented highlights of the organization and statistical data about WAM, the high school lectureship sponsored by the MAA under a grant from IBM since 1975. The discussion centered on the results of the regular evaluations of WAM visits by students, host teachers, and guidance counselors as well as on the pilot survey conducted in Spring, 1977, in seven New Jersey schools to study attitudes toward mathematics. A perspective of the future of the WAM program was described.

Professor Pedersen presented a brief anecdotal, and informal, account demonstrating some of the similarities and differences between WAM and other lectureship programs sponsored by the MAA.

Mathematics and Mathematicians In Industry, by Dr. Brockway McMillan, Vice President, Military Systems, Bell Telephone Laboratories.

The role of the mathematician as a consultant in an industrial setting was examined. Some implications were pointed out, bearing on his education, and upon his attitudes as well as upon those of his employer.

THIRD SESSION OF THE ASSOCIATION

Hedrick Lecture III: Binocular Vision and Mach Bands, by Professor Joseph B. Keller.

Some of the early experimental results concerning binocular vision were described. Then the notion of visual space was explained, and Luneburg's introduction of a non-euclidean metric into that space was presented to account for those results. Blank's modification of the metric to fit later experimental results was also described. The question of uniqueness of the metric was considered too. Applications of the theory, which involve hyperbolic geometry, were touched upon.

Mach's experiments, leading to the observation of his famous bands, were described. A mathematical theory which accounts for these bands was explained. This theory was then related to the observations of lateral inhibition in the visual system of the horseshoe crab.

Business Meeting of the Association: Presentation of Carl B. Allendoerfer, Lester R. Ford, and George Pólya Awards, announcement of a bequest and gift to the Association.

On the Landau Problem of Bounds for Derivatives, by Professor Isaac J. Schoenberg, United States Military Academy, West Point.

Let Γ be a smooth curve in the complex plane, and let $\alpha(t)$ $(-\infty<t<\infty)$ be a "motion" on Γ, i.e., $\alpha(t) \in \Gamma$ for all t, having the velocity vector $\dot{\alpha}(t)$, and acceleration vector $\ddot{\alpha}(t)$. Let $A>0$ be given. It is shown that among all the motions on Γ, such that (1) $|\ddot{\alpha}(t)| \leq A$ for all t, there is a unique motion $\tilde{\alpha}(t)$, called the Landau motion on Γ, for the constant A, such that at every point P of Γ, the speed $|\dot{\tilde{\alpha}}(t)|$ is $\geq$ the speed $|\dot{\alpha}(t)|$ at P, of any other motion $\alpha(t)$ satisfying (1). The Landau motion $\tilde{\alpha}(t)$ is explicitly determined for a few simple cases, such as if Γ is a circle, an arc of a circle, a parabola, or a cycloid.

FOURTH SESSION OF THE ASSOCIATION

Measure Algebras and Their Uses, by Professor Dorothy Maharam Stone, University of Rochester.

All measures considered are σ-finite. This survey included the following topics: The algebras; the representation of a measure algebra with respect to a given subalgebra; the realization of automorphisms and homomorphisms by point mappings; applications to lifting theory, to disintegration of measure spaces, to the structure of function spaces, and to positive operators on function spaces.

One Step Beyond: Some Well-Known Results We Seldom See, by Professor Bill Leonard, California State University, Fullerton.

A mathematics professor once remarked, "I was a junior in college before I knew that a math major studied anything other than calculus." It can be argued that four semesters of calculus often

leave a math student less than culturally satisfied. This talk suggested some ancillary topics which, while basically in the field of elementary calculus, are seldom presented. These included some notes on the number e, a brief look at a fascinating result due to Euler, and a hint at a topic in complex analysis.

Some Recent Results in the Geometry of N-space, by Professor David G. Larman, University of Washington and University College, London.

Say that a set S in Euclidean n-space, E^n, realizes a distance d if there exists two points x and y in S whose distance apart is d. Considerable progress has been made in recent years on two longstanding problems concerning the realization of distances. The first is to determine the least number f(n) such that if a set in E^n realizes only two distances then its cardinality is at most f(n). The second is to determine the largest number g(n) such that if g(n) sets cover E^n then at least one set realizes all distances. It is now known that $f(n) \leqq 1/2(n+1)(n+4)$ and $\lim_{n\to\infty} g(n)/n^{\alpha}=\infty$ for any definite α.

MAA-NCTM Sourcebook of Applications of Secondary Mathematics, by Professor Donald W. Bushaw, Washington State University.

For several years a committee sponsored jointly by MAA (through CUPM) and NCTM, with NSF support, has been preparing a sourcebook of applications of secondary school mathematics. This has been in response to a long-felt need for instructional materials based on authentic applications of standard topics in the pre-calculus curriculum, many of which are not shown to be applicable at all in most existing textbooks. The project is now near completion, and this talk was a report on it, with emphasis on possible uses of the sourcebook in two-year colleges.

A Lost Notebook of Ramanujan, by Professor George E. Andrews, Pennsylvania State University.

Ramanujan wrote a further notebook apparently during the last year of his life. A discussion of its origin was given. The notebook contained over 600 formulas listed without proof. About 400 are related to the theory of q-series. Several of these results were discussed together with possible applications in number theory.

Creative Gems from the Classroom, by Professor Robert J. Bitts, Arapahoe Community College.

A few small gems from students were offered as evidence that teachers, students, and even math in the two-year colleges are okay. The "gems" ranged from the Brothers Comma, as in 0.001,23 to the sublime Spanish mnemonic for π given to the author by a Salvadorean village lad, Chico Herrera: "Sol y Luna y Mundo proclaman al eterno Autor del Cosmos!"

The Place of Unsolved Problems in the Classroom, by Professor Ben J. Jones, Green River Community College.

The author believes that much of the lack of interest in mathematics encountered in the classroom is due to the bad press mathematics receives from society at large; it has the reputation of being "cut-and-dried,mechanical, and it's all known."

In order to combat this, the author included in his courses unsolved and unsolvable problems (a distinction which students tend to find philosophically fascinating). This brief talk discussed two such problems appropriate to different levels.

1) The Twin Primes conjecture, and Brun's Theorem.

2) Finite integration, from elliptic integrals to a recent theorem on the lack of an algorithm telling if a given function can be integrated in closed form.

SPECIAL SESSIONS OF THE ASSOCIATION

Film showings were held in Room 130, Kane Hall, on Sunday and Monday evenings at 7:00 P.M. The following films were shown:

Sunday

7:00-7:09 P.M.	Dot and the Line
7:10-7:26 P.M.	Curves of Constant Width (with J.D.E. Konhauser)
7:30-7:55 P.M.	Time for Change—The Calculus (a BBC broadcast as part of the Open University's History of Mathematics)
8:00-8:11 P.M.	Mathematics Peep Show
8:15-8:23 P.M.	Similar Triangles
8:25-8:37 P.M.	Flatland

8:40-8:50 P.M.	The Theorem of the Mean-a film of the MAA Calculus Film Series
8:52-9:00 P.M.	Newton's Equal Areas
9:03-9:25 P.M.	Cycloidal Curves or Tales from the Wanklenberg Woods

Monday

7:00-7:34 P.M.	Films of the Topology Films Project
7:00-7:14 P.M.	Regular Homotopies in the Plane: Part I
7:15-7:34 P.M.	Regular Homotopies in the Plane: Part II
7:35-8:21 P.M.	The Marriage Theorem: Applications, Part II (b&w)
8:25-8:38 P.M.	Dihedral Kaleidoscopes (with H.S.M. Coxeter)
8:40-8:57 P.M.	Films Produced by Thomas F. Banchoff and Charles Strauss
8:40-8:47 P.M.	The Hypercube-Projections and Slicing
8:50-8:57 P.M.	Complex Functions Graphed in 4-Space (b&w)
9:00-9:10 P.M.	Newton's Method (MAA Calculus Film Series)
9:12-9:20 P.M.	Accidental Nuclear War (produced by David Gillman)

MEETING OF THE BOARD OF GOVERNORS

The Board of Governors met on Saturday at 9:00 A.M. in the Condon Room of the University Tower Hotel with forty members present. Among the items of business transacted were the following:

The Board elected Professor Leonard Gillman of the University of Texas at Austin as Treasurer for the period 1978-82. The Board also elected Professor Donald J. Albers as Editor of the TWO-YEAR COLLEGE MATHEMATICS JOURNAL for the period 1979-83.

As Associate Editors of the AMERICAN MATHEMATICAL MONTHLY, for the period extending through 1981, the Board elected Professors Paul R. Halmos, University of California, Santa Barbara, Paul A. Haeder, University of Nebraska at Omaha, and Timothy J. Robertson, University of Iowa. Professor Halmos will edit a new section of the MONTHLY entitled *Progress Reports*. Professor Haeder will edit *News and Notices* and Professor Robertson will serve as a consultant on statistics.

Finally, the Board elected Professor Stanley Friedlander, Bronx Community College, and Professor Paul Campbell, Beloit College, as Associate Editors of the TWO-YEAR COLLEGE MATHEMATICS JOURNAL and MATHEMATICS MAGAZINE, respectively. Professor Friedlander will serve as Assistant Editor of the *Problems and Solutions* section of the TYCMJ for the period extending through 1978 and Professor Campbell will edit the *Reviews* section of MATHEMATICS MAGAZINE during the period extending through 1980.

The Board of Governors voted to receive the following grants:

a. A grant of $5850 from the Office of Naval Research for support of the 1977 Olympiad Training Session.
b. A grant of $4400 from NSF to complete the work of producing and distributing "Modules in Applied Mathematics" produced at the 1976 Summer Workshop at Cornell University.
c. A grant of $10,000 from IBM for support of the secondary school lectureship program "Women and Mathematics" during the 1977-78 academic year.

The Board approved the following schedule of meetings:

Sixtieth Summer Meeting: University of Oregon, August 18-20, 1980

Sixty-second Annual Meeting: Biloxi, Mississippi, January 26-28, 1979

The second of these replaces Milwaukee, Wisconsin, which had previously been approved as the site for the 1979 Annual Meeting. With these additions, the meetings authorized are:

Atlanta, Georgia	January 6-8, 1978
Brown University	August 8-10, 1978
Biloxi, Mississippi	January 26-28, 1979
Virginia Polytechnic Institute & State University	August 21-23, 1979
San Antonio, Texas	January 5-7, 1980

University of Oregon	August 18-20, 1980
San Francisco, California	January 10-12, 1981

The Board approved a contribution of $5000 toward support of a Congressional Science Fellow. Both SIAM and the AMS are cooperating with the Association in support of this program. Under this program, a mathematician will be assigned to a Congressional staff position for a one-year period. The program is to be modeled after the AAAS program and will be implemented with that organization's cooperation. A detailed announcement of the Congressional Science Fellowship will appear in an early issue of the MONTHLY. Both new Ph.D.'s and persons on sabbatical are encouraged to apply.

The Board approved a *CEEB/MAA Statement on the CLEP Examination.* This Statement was prepared by the CEEB/MAA Committee on Mutual Concerns in an attempt to meet the concerns expressed by MAA members about misuses of the CLEP (College Level Examination Program) by some institutions. It is hoped that the College Entrance Examination Board will also endorse this Statement.

Leonard Gillman, Treasurer, reported to the Board that the Association showed an excess of income over expenditures of $241,000 for 1976. This is because of three special gifts, totalling $170,000, to the Association. The remaining balance of $71,000 just about offsets the deficit of $77,000 in 1973.

The Board approved a statement entitled *On the Preparation Needed by Students Planning to Take Collegiate Mathematics.* This statement, prepared by a joint MAA-NCTM Committee, chaired by Professor Gerald L. Alexanderson, will be submitted to the NCTM Board of Directors for consideration at their September, 1977, meeting.

Professor Malcolm W. Pownall reported to the Board that, during the 1976-77 academic year, 97 colleges arranged for visits by lecturers through the Committee on Visiting Lecturers and Consultants. Also, 7 colleges were visited by consultants. Professor Pownall reported that there had been considerable demand for the non-academically employed lecturers.

The Board discussed at some length the possibility of the Association's purchasing a headquarters building. There are now on hand two large gifts for this purpose. Sites in Washington, D.C. are currently being investigated for possible purchase.

The inauguration of a subscription program to enable departments and/or colleges to participate in an MAA Placement Testing Program was announced. The initial fee for this program is $75 and an announcement of the Program has been mailed to all department chairmen. This Program makes available three algebra tests and a trigonometry/elementary functions test. All of the tests have been used in a two-year pre-testing period and have been developed with the cooperation of CEEB and the Educational Testing Service.

President Alder reported to the Board on the highly favorable reaction to the brochure, *The Math in High School . . . You'll Need for College,* produced in cooperation with NCTM. This brochure was mailed to all Association members and to a large selection of guidance counselors. MAA members are encouraged to call this brochure to the attention of secondary school students, administrators, and faculty.

BUSINESS MEETING OF THE ASSOCIATION

A business meeting was held at 10:00 A.M. on Monday in Meany Hall with President Alder presiding.

President Alder announced certain changes in the administration of the awards for expository articles in Association journals. Three Awards are now to be presented. These are the Carl B. Allendoerfer Awards for articles in MATHEMATICS MAGAZINE, the George Pólya Awards for articles in the TWO-YEAR COLLEGE MATHEMATICS JOURNAL, and the Lester R. Ford Awards for articles in the MONTHLY. At most, two Allendoerfer and Pólya Awards are to be made annually and, at most, five Ford Awards are to be made annually (for further details on these Awards, see the August-September issue of the MONTHLY).

Prior to announcing the Allendoerfer Awards, President Alder called on Professor C. O. Oakley, who addressed the business meeting as follows:

"It is not our purpose at this time to review in detail the life and works of Carl Allendoerfer who died 29 September 1974. Since some of our members may not have known him personally I should like to mention some references. One can see something of the man and his accomplishments in the announcement of the "Award for Distinguished Service to Professor Carl Barnett Allendoerfer" in the MONTHLY, Vol. 79(1972) and in the remarks of Professor B. W. Jones, made at the Annual Business Meeting of the Association, 26 January 1975. I would also call to your special attention the very sensitive article on Carl written by his Oxford classmate, F. A. Ficken. You will find it in The American Oxonian, Vol. LXI, No. 4, 1974.

Carl's life was a life of complete dedication to mathematics and he spent time, energy and capital in trying to help the cause of mathematics in general and the works of the Association in particular. He moved quietly and efficiently and often anonymously. For example it was his agreement with the Association that only after his death were we to be told that the Greenwood Fund (some $25,000) was a gift of his to be used in film-making.

Another bequest of the Allendoerfers, made by Carl, in his will, and added to later by his wife, Dorothy, is in the amount of $50,000. With the income from this fund the Association has established the Carl B. Allendoerfer Awards. These awards are for expository articles appearing in the Mathematics Magazine and the first of them are to be presented by Mrs. Allendoerfer herself in just a few moments.

Our Association is deeply indebted to the Allendoerfers. For their generous gifts and for the many things that Carl did for the Association during his years of unstinted service, we are indeed grateful and we extend our sincere thanks and best wishes to Dorothy.

I should like to share with you the last words Carl and I were to speak to each other. They attest both to his indomitable strength of character and to his good humor. Carl said: 'Cletus, go to my office and if you see things there that you'd like to have, take them with you. Where I'm going, I won't need them.' A smile flickered across his pain-racked face. I countered with the question: 'Well, Carl, have you been practicing up on your harp?' 'Yes,' he said, 'I have.' 'But just in case, I'm also finding out the best ways of putting out fires!'

Throughout life and into death he never broke his stride."

Mrs. Dorothy H. Allendoerfer then presented the first set of Carl B. Allendoerfer Awards for articles published in 1976 written by Professors Joseph A. Gallian and B. L. van der Waerden. Professor Gallian was present to accept the Award.

Professor George Pólya presented the first set of George Pólya Awards for articles published in 1976 to Professors Anneli Lax and Julian Weissglass.

President Alder presented the twelfth set of Lester R. Ford Awards for articles published in 1976. Present to accept these Awards were Professors Shreeram S. Abhyankar, William H. Gustafson, Paul R. Halmos, James P. Jones, Joseph B. Keller, Daihachiro Sato, and Hideo Wada.

President Alder announced that the team of eight secondary school students from the USA which competed in the 1977 International Mathematical Olympiad (IMO) in Belgrade, Yugoslavia, had scored more points than any other team from the twenty-one countries entered in the competition. He said that this was a remarkable achievement and that the Association had received many messages of congratulations on the success of the team. In particular, he read a letter of congratulations from President Jimmy Carter.

President Alder introduced Professor Nura D. Turner, Advisor to the Olympiad Awards Ceremony, and Professors Samuel L. Greitzer and Murray S. Klamkin, the coaches since the first competition by a USA team in the IMO in 1974.

Professor Greitzer gave a brief description of how an IMO is conducted. He included mention of the Olympiad Training Session held in 1977 at West Point and foreign travel in addition to details of the IMO. The latter details included how questions are selected and translated, how the examination is administered, and how the papers are corrected. Professor Greitzer also told of the recreational and sight-seeing activities provided participants by the host country. He also announced that the 1978 IMO will be held in Rumania.

The Secretary presented his report announcing first some of the actions of the Board of Governors and the financial report for 1976. The Secretary also announced that on July 1, 1977, there were 17,174 individual members in good standing compared to 17,174 individual members in good standing on July 1, 1976. The number of Academic Members on July 1, 1977, was 424 compared to 407 a year earlier. The Secretary noted that the COMBINED MEMBERSHIP LIST includes a list of Academic Members of the Association. He suggested that persons present determine whether their institution is an Academic Member of the Association and, if not, that they inquire about joining.

The Secretary reported that the possibility of inviting the IMO to be held in this country was now being investigated. He said that the Association was indebted to the Army Research Office, IBM, and the Office of Naval Research for their continuing support of the Olympiad activities.

The Secretary expressed thanks to the members of the Committee on Arrangements, especially to Professor Roy Dubisch for having chaired the Committee. The Secretary also called attention to the fine program for the meeting and thanked the members of the Program Committee: Victor L. Klee, *Chairman;* Brian R. Alspach, Ross A. Beaumont, Hugh D. Brunk, Donald W. Bushaw, Richard K. Guy, Joseph Hashisaki, Jean J. Pedersen, Howard E. Zink.

MEETING OF SECTION OFFICERS

The meeting of representatives of the Sections was held on Sunday, August 14, 1977, in the Condon Room of the University Tower Hotel. Dean Lester H. Lange, Chairman of the Committee on Sections, presided. Sixty-four persons were present, representing twenty-eight of the twenty-nine Sections. Members of the Committee on Sections present at the meeting were Professors James C. Bradford, Louis A. Guillou, Samuel W. Hahn, Jacqueline C. Moss, and Alfred B. Willcox. Peggy Perdue served as recorder.

President Henry L. Alder first addressed the Section Officers as follows:

"It is a great pleasure to welcome all of you to this meeting of Section Officers. As, I am sure, you are all aware, this meeting serves an extremely important role, namely to give you, the Section Officers, an opportunity to learn first hand of activities, projects and other items which you can adopt in your Section in order to increase its effectiveness, that is, to carry out in the best possible way the goals of the Association within your region.

I hope very much that you will find this meeting profitable and are able to carry many good ideas back to your Section for implementation.

Many Sections already carry out the goals of the Association in an exemplary fashion. I have been thoroughly impressed with the organization, administration and programs of several of the seven Sections I had the privilege of visiting so far this year. In fact, I liked several projects and activities so well that I suggested to the Chairman of the Committee on Sections that they be included in the program of the meeting tonight. As a result, they will be described later in this program and, therefore, there is no need for me to mention them.

Finally, let me urge all of you to let us know of anything we, as national officers can do, to help you with some of the problems you may encounter in your Section. If you don't find the opportunity to mention them here tonight, by all means talk to me after the meeting or write to me. I want you to know that all your comments will be taken very seriously."

Dean Lange called attention to the folder of MAA information suitable for passing from a Section Secretary to his or her successor. This folder is being assembled by the Committee on Sections and will be available during 1978.

Dean Lange also invited suggestions for a revision of the booklet, *Guidelines for Sections*. Professor Louis A. Guillou of the Committee on Sections is now preparing a revision.

Next, Dean Lange called on the Section Officers to recommend persons to President Alder for service on MAA committees. He said that persons who have been active in Section activities would be especially appropriate for such appointments.

The meeting next turned to the first agenda item; how to improve attendance at Section meetings. Leonard Gillman, MAA Treasurer, said that he had visited several Section meetings and found that some were more active than others. He conjectured that the most successful meetings were those attended by a mix of university, four-year college, and two-year college faculty. He also pointed to the apparent success of meetings at which the invited speakers were well-known. Another factor he identified was special sessions such as, for example, student papers, sessions on calculators, etc.

Professor Alavi, Michigan Section, reported that the Michigan Section had tripled its attendance by beginning two-day meetings with a cocktail party and after-dinner speaker on the evening between the sessions. Professor Alavi also said that the Michigan Section has utilized numerous speakers from industry with considerable success.

Professor James C. Bradford, Committee on Sections, reported on attendance at Section meetings. He said that attendance figures reported in the AMERICAN MATHEMATICAL MONTHLY for the past three years show the largest attendance to be 274 and the lowest to be 44. In terms of percentages, attendance at Section meetings has ranged from 61% to 6% of the members in the Sections.

Professor B. E. Gillam, Drake University, reported that it was difficult to get MAA members from the large state universities to attend Section meetings. President Alder suggested that it might be possible to involve persons from the larger schools to serve as officers. He cautioned against ruling out such persons as officers on the premise that they would probably refuse.

It was reported by Professor David Ballew that, in 1980, the Rocky Mountain Section of the MAA would hold a joint meeting with the Pacific Region of the AMS. Professor Lida Barrett replied that such joint meetings had not contributed greatly to the success of the meetings of the Southeastern Section. Consequently, joint AMS-MAA meetings are now being discontinued by that Section. Professor I. C. Gentry indicated that the Southeastern Section has strong support from its major universities. He said that the meetings of the Southeastern Section have excellent programs and are well attended except possibly when the meeting is held near one of the boundaries of that Section.

It was reported by Professor Ellen C. Veed that the Kansas Section holds joint meetings with the Kansas Association of Teachers of Mathematics. Professor William H. Beyer reported that the Ohio Section holds two-day meetings with the evening between used for swap sessions and panel discussions. Professor Barrett indicated that sessions for contributed papers by students had been successful in the Southeastern Section.

The activities of the Oklahoma-Arkansas Section were reviewed by Professor E. K. McLachlan. He pointed out that the ingredients necessary for success of the Section's activities include a good Section Secretary and the strong support of the chairmen of the mathematics departments at institutions within the Section.

Professor Gary Sherman said that it seemed necessary that there be one person in each department who would encourage department members to attend Section meetings. He said that the department chairman would be a suitable person. Newman H. Fisher suggested that the MAA departmental representative might also be useful in this role.

Professor Fisher also reported on the activities of the Northern California Section. In particular, he mentioned a luncheon program on math anxiety and a meeting of departmental chairmen following the meeting. The Secretary noted that the Texas organization of administrators in the mathematical sciences meets in conjunction with the meeting of the Texas Section.

The difficulties caused by the geography of Sections were commented on by Ernest C. Schlesinger and Robert A. Chaffer. Professor Chaffer pointed out that while the Michigan Section does not hold meetings in the northern parts of its Section, there have been summer institutes held in that area.

Professor George W. Lofquist, Eckerd College, said that the Florida Section usually meets in conjunction with Pi Mu Epsilon. He noted that the host institution is required to meet the expenses of a visiting speaker. Professor Lofquist also reported on the mini-Sectional meetings held in Florida during the past year. These meetings covered small geographic areas and were designed to increase communication between the Florida Section and two-year college and secondary school faculty. These meetings were held on a single day and the talks given were, for the main part, pedagogical in nature. Professor Charles W. McArthur noted that about 300 persons had attended mini-Section meetings. He pointed out that the small geographic area allowed persons involved to concentrate on turning out the local secondary school and two-year college faculty.

Robert L. Wilson encouraged MAA Sections to include talks by members of Mu Alpha Theta, a secondary school mathematics club sponsored by the MAA, on the programs of their meetings. He said that Sections interested in having a list of Mu Alpha Theta chapters in their area should write to Professor Harold V. Huneke, Mathematics Department, University of Oklahoma, Norman, OK. 73019.

Dr. Henry O. Pollak pointed out that attendance at MAA Section meetings may be related to the overall mathematical activity within an area. In areas where the Section meeting is the major activity, the meetings may be better attended than in areas where there is a great deal of other activity available. He felt that inclusion of papers by students, both secondary and college, might improve participation by college and university faculty in the Section meetings.

Alfred B. Willcox, MAA Executive Director, recommended that Governors organize meetings with MAA departmental representatives at Section meetings.

Professor J. Thomas Renfrow, Beloit College, said he felt that, if a Section planned consistently good programs, then attendance would eventually increase. He also advocated that the Sections include students in their programs and that individual faculty members transport their students to the Section meetings.

Professor Samuel W. Hahn spoke on the success of the Ohio Section Newsletter. He said it was an inexpensive way to get a great deal of information to members of the Section. This led to a discussion of newsletters during which it was said that newsletters may be sent by bulk mail and are financed by dues and other income. Professor Hahn said that it costs the Ohio Section less to duplicate and mail three eight-page issues of its newsletter than it previously cost to send the meeting announcements by first-class mail. Professor Steve Galovich said that the North Central Section was planning to ask for donations for support of their newsletter. Professor Alavi said the Michigan Section supported its newsletter by means of academic memberships. It was noted that each Section Secretary receives the newsletters from all Sections.

Professor James C. Bradford, Committee on Sections, reported the results of a questionnaire on Section attendance. The questionnaire was sent to two large Sections and two small Sections; one of each category with good attendance and one of each category with poorer attendance. An interesting result of this survey was that 62% of the respondents from the Sections with good attendance indicated that informal visits with colleagues was an important reason for them to attend Section meetings. Only 38% of the respondents from the less well-attended Sections indicated that this was important. Professor Bradford said that it was very encouraging that only 13% of the respondents indicated that the programs at Section meetings were not of sufficient interest to attend. He also emphasized that a Section should pay attention to personal items such as retirements, awards, etc.

President Alder stressed the importance of careful organization of the Section meetings. He said great care should be taken in the selection of Section officers, particularly the Secretary-Treasurer, in order to ensure carefully organized meetings.

In response to a question, Dean Lange said that the Committee on Sections had not studied a reorganization of the boundaries of the Sections. He pointed out that the impetus for changes of this kind has come from the Sections rather than the national organization.

Dean Lange said that Section Chairmen are to be asked to bring copies of the programs for the 1977-78 Section meetings to the 1978 meeting of Section Officers at Brown University. These are to be used as a reservoir of ideas for Section programs.

Professor Jane Day said that the Northern California Section plans to make use of the brochure *Math in High School*. She said that it was hoped that this could be distributed to each ninth grader in the Bay area.

Professor Louis A. Guillou reported on the efforts of the Committee on Sections to prepare a revision of *Guidelines for Sections*. A review of the type of information being included in the pass-the-baton type of booklet being prepared for Section Officers was presented and additional suggestions were solicited. Professor Lily Christ suggested a section on the legal responsibilities and guidelines for each Section and inclusion of the dates when certain information is available or required; e.g., the results of the Annual High School Mathematics Examination. Professor David Ballew suggested that there should be included some successful program items.

Professor Guillou reported that he had surveyed the Sections on dues and had forwarded the responses to all Section Secretaries. Anyone else wanting this information should write to Professor Guillou.

DINNER IN HONOR OF THIRTY-YEAR MEMBERS

On Monday, the Association held a dinner in honor of its thirty-year members. This dinner was held at 6:30 P.M. in the Ballroom of the University Tower Hotel. A similar dinner was held at the August, 1976, meeting in Toronto. Like the earlier dinner, this was designed as a pleasant occasion to recall the services of the MAA's senior members and to inform them of current activities and future plans. Professor Raymond L. Wilder served as toastmaster and President Henry L. Alder brought greetings from the MAA. The speakers were Professor Edwin F. Beckenbach, *The MAA Publications Program;* Professor Victor L. Klee, Jr., *Some Presidents I Have Known;* and Professor George Pólya, *Epilogue*.

One hundred seven persons attended the dinner.

MEETINGS OF OTHER ORGANIZATIONS

The Pi Mu Epsilon Fraternity held its sessions for contributed papers on Monday at 3:00 P.M. and on Tuesday at 3:00 P.M. The Pi Mu Epsilon Governing Council Luncheon was held at noon on Monday in the Husky Den of the HUB. The banquet was held at 6:00 P.M. in a local restaurant and the J. Sutherland Frame Lecture was delivered at 8:00 P.M. on Monday in Room 110 of Kane Hall. The lecture was entitled *Techniques for Solving Extremal Problems* and was delivered by Professor Ivan Niven of the University of Oregon.

The Association for Women in Mathematics held a panel discussion entitled *Alternatives to Academic Employment for Mathematicians*. This discussion was held at 7:30 P.M. on Monday in Room 120 of Kane Hall and Professor Lenore Blum served as moderator. At 4:00 P.M. on the same day, the AWM Executive Committee held an open meeting in Room 409 of Guggenheim Hall.

The Council of the Conference Board of the Mathematical Sciences met on Wednesday at 2:15 P.M. in the Condon Room of the University Tower. The Mathematicians Action Group held an open meeting of its Steering Committee at 7:30 P.M. on Sunday, its Business Meeting at 3:15 P.M. on Monday, and a panel discussion at 9:30 A.M. on Tuesday. These meetings were held in Room 110 of Kane Hall.

The Institute of Mathematical Statistics held sessions from Monday, August 15, through Thursday, August 18. There were scheduled more than twenty invited papers and several contributed paper sessions for Room 110 and 120, Kane Hall. In addition, the IMS Rietz Lecture was delivered at 11:00 A.M. on Monday by Bradley Efron and the Wald Lectures were delivered at 2:15 P.M. on Monday, 1:30 P.M. on Tuesday, and 2:15 P.M. on Wednesday. Both lectures were scheduled for 120 Kane Hall. The IMS Business Meeting was held at 5:00 P.M. on Tuesday in 120 Kane Hall.

The American Mathematical Society held sessions from Monday, August 15, through Thursday, August 18.

The Data Subcommittee of the AMS Committee on Employment and Educational Policy sponsored an open meeting and panel discussion at 3:30 P.M. on Monday in Meany Hall. The panelists included Professors Lida K. Barrett and Wendell H. Fleming.

All invited AMS lectures were given in Meany Hall according to the following schedule:

Monday

1:00 P.M. AMS COLLOQUIUM LECTURE: *Geometric Measure Theory*, Professor Herbert Federer, Brown University.

Tuesday

8:30 A.M. *Pattern Formation and Periodic Structures in Systems Modeled by Reaction-Diffusion Equations*, Professor James M. Greenberg, SUNY at Buffalo.

9:45 A.M. *Symmetric Structures in Banach Spaces*, William B. Johnson, Hebrew University, Jerusalem, and Ohio State University, Columbus.

11:00 A.M. COLLOQUIUM LECTURE II: Professor Herbert Federer

Wednesday

10:00 A.M. AMS Prize Session

11:00 A.M. AMS Business Meeting

1:00 P.M. COLLOQUIUM LECTURE III: Professor Herbert Federer

Thursday

8:30 A.M. *Interplay Between Classical Modular Forms and Associated Galois Representations*, Kenneth A. Ribet, Princeton University.

9:45 A.M. *What is a Topological Manifold? (The Characterization Problem)*, James W. Cannon, University of Wisconsin, Madison.

11:00 A.M. *Some Aspects of the Theory of Elliptic Equations in Differential Geometry*, Shing-Tung Yau, Stanford University.

1:00 P.M. COLLOQUIUM LECTURE IV:
Professor Herbert Federer

ARRANGEMENTS, ENTERTAINMENT, AND RECREATION

The Committee on Arrangements consisted of Roy Dubisch, *Chairman;* Kathleen Baxter, Ross A. Beaumont, Samuel L. Dunn, Thomas W. Hungerford, J. Maurice Kingston, Norman C. Meyer, Jr., Lloyd J. Montzingo,Jr., David P. Roselle, Kenneth A. Ross, Fredrick W. Scholz, Gordon L. Walker.

Registration headquarters were located in the Basement Registration Area of Odegaard Undergraduate Library. Books and educational material were displayed in the registration area on Sunday, Monday, and Tuesday. The Mathematical Sciences Employment Register was conducted on an informal basis with facilities to display resumes and job listings.

An Indian-style barbecue was served at 6:00 P.M. on Tuesday. A beer party followed the barbecue at 8:00 P.M. and was held in the Ballroom of the Student Union Building.

David P. Roselle, *Secretary*

ACKNOWLEDGEMENT

The editors wish to thank the following individuals who have refereed manuscripts for Volume 84: Harvey L. Abbott, János D. Aczél, Ronald Alter, P. M. E. Altham, R. D. Anderson, Henry A. Antosiewicz, Thomas M. Apostol, Richard Askey, Krishna Athreya, R. G. Ayoub, George Bachman, Robert G. Bartle, Jon K. Barwise, Steven F. Bauman, Gilbert Baumslag, John V. Baxley, Georgia Benkart, Claude Berge, Elwyn Berlekamp, Bruce Berndt, Frank Bernhart, Leon Bernstein, Lipman Bers, Richard Bieberich, Ian Blake, Michael N. Bleicher, Mary Boas, R. P. Boas, Thomas K. Boehme, William E. Boyce, Fred Brauer, George U. Brauer, Leon Breiman, Joel Brenner, John D. Brillhart, Felix Browder, Ezra Brown, W. Dale Brownawell, A. M. Bruckner, B. F. Bryant, Thomas H. Brylawski, R. Creighton Buck, F. R. Buianouckas, H. S. Butts, R. N. Cain, J. Calvert, James Cannon, Gerald Cargo, Leonard Carlitz, David Carlson, Gulbank Chakerian, Donald R. Chalice, Gary Chartrand, Earl Coddington, Howard Conner, Kenneth L. Cooke, R. L. Cooley, H. S. M. Coxeter, Donald W. Crowe, Frederic Cunningham, Jr., C. W. Curtis, Richard Darst, Chandler Davis, Richard A. Dean, Carl-Wilhelm R. De Boor, Emeric Deutsch, Allen Devinatz, Harold Diamond, R. W. Dickey, Raymond F. Dickman, Larry L. Dornhoff, James Dugundji, Peter L. Duren, Meyer Dwass, Verena H. Dyson, Paul Eakin, Roger B. Eggleton, J. A. Eidswick, David Eisenbud, Arthur Erdélyi, Ronald J. Evans, Martha Evens, Edward R. Fadell, L. Fejes-Tóth, P. A. Fillmore, Stephen Fisher, Harley Flanders, Frank Forelli, Aviezri Fraenkel, J. Sutherland Frame, Wolfgang H. Fuchs, Adriano M. Garsia, Peter M. Gibson, Andrew M. Gleason, Casper Goffman, R. R. Goldberg, Seymour Goldberg, Solomon W. Golomb, Henry W. Gould, Colin C. Graham, Curtis Greene, Priscilla Greenwood, L. Gross, Branko Grünbaum, Stanley P. Gudder, Hiroshi Gunji, William H. Gustafson, Paul R. Halmos, Kenneth B. Hannsgen, Bernard Harris, Simon Hellerstein, Edwin Hewitt, Donald Higman, Paul Hill, J. G. Hocking, Richard E. Hodel, R. A. Honsberger, Frank N. Huggins, Martin I. Isaacs, James A. Jenkins, Arnold Johnson, F. Burton Jones, B. Jónsson, James Jordan, Wolfgang B. Jurkat, D. Kahn, David Kay, Nicholas D. Kazarinoff, H. Jerome Keisler, Murray Klamkin, Daniel J. Kleitman, David A. Klerner, Morris Kline, Bernard Kolman, James D. Kuelbs, Kenneth Kunen, R. A. Kunz, Thomas G. Kurtz, J. Lamperti, Eric S. Langford, Edward Larsen, Joseph LaSalle, J. D. Lawson, Solomon Leader, Derrick Lehmer, Emma Lehmer, H. W. Lenstra, Lawrence Levy, W. Lindstrom, T. P. Liverman, W. F. Lucas, W. F. Lunnon, Donald A. Lutz, David P. Maher, K. Mahler, Joseph Malkevitch, A. Manaster, Michael Marcus, Martin J. Marsden, Kenneth May, Russell L. Merris, Paul T. Mielke, John W. Milnor, Leon Mirsky, Cleve Moler, A. R. Møller, D. Moore, A. P. Morse, K. R. Mount, James A. Murtha, R. A. McCoy, Daniel R. McMillan, Alex Nagel, Zeev Nehari, Morris Newman, Peter Ney, W. Keith Nicholson, Harald Niederreiter, I. Niven, William P. Novinger, Andrew M. Odlyzko, C. D. Olds, Ingram Olkin, F. W. J. Olver, T. G. Ostrom, James C. Owings, Jr., John C. Oxtoby, E. W. Packel, Emanuel Parzen, Donald Passman, Stanley E. Payne, Joan Wick Pelletier, Martin Pettet, Robert R. Phelps, Mark Pinsky, G. Piranian, Carl Pomerance, Jack R. Porter, M. Pownall, J. J. Price, Burton Randol, William L. Reddy, Ray Redheffer, Kenneth B. Reid, Irma Reiner, Stanley Reiter, James Retherford, W. C. Rheinboldt, Leonard Richardson, H. Richert, Jerry R. Ridenhour, G. Ringel, F. S. Roberts, T. J. Robertson, C. A. Rogers, D. P. Roselle, Edward B. Saff, G. T. Sallee, Richard Savage, Hans Schneider, I. J. Schoenberg, Daniel R. Scholz, Seymour Schuster, Abraham Schwartz, Richard A. Scoville, Daniel Shea, Paul C. Shields, James E. Shockley, David Singmaster, Richard Sinkhorn, David A. Smith, Louis Solomon, Joel Spencer, Ivar Stakgold, Richard Stanley, Sherman Stein, Frank Stenger, S. Stigler, Gilbert W. Strang, Karl R. Stromberg, Wm. C. Swift, Richard A. Tapia, O. Taussky-Todd, Jean E. Taylor, John G. Thompson, Robert C. Thompson, William T. Trotter, Alan Tucker, Jacobus H. Van Lint, Stephen Wainger, Seth Warner, James K. Washenberger, Daniel Waterman, Robert J. Weber, Guido L. Weiss, James H. Wells, James G. Wendel, Robert J. Whitley, Sylvia Wiegand, David E. Wilson, John S. Wilson, J. Wolfowitz, Gordon Woodward, Theophil Worosz, E. M. Wright, D. Ylvisaker, Gail S. Young, Lawrence A. Zalcman, Daniel Zelinsky, Phillip Zenor.

CALENDAR OF FUTURE MEETINGS

Sixty-first Annual Meeting, Atlanta, Georgia, January 6–8, 1978.

Fifty-eighth Summer Meeting, Brown University, August 8–10, 1978.

The following is a list of the Sections of the Association with dates of future meetings so far as they have been reported to the Editorial Director.

ALLEGHENY MOUNTAIN, University of Pittsburgh, Pennsylvania, April 14–15, 1978.

FLORIDA, St. Petersburg Junior College, Clearwater, March 3–4, 1978.

ILLINOIS, Western Illinois University, Macomb, May 5–6, 1978.

INDIANA, Earlham College, Richmond, April 22, 1978.

INTERMOUNTAIN

IOWA, University of Northern Iowa, Iowa Falls, April 22, 1978.

KANSAS, Wichita State University, Wichita, late March-early April 1978.

KENTUCKY, early April. Deadline for papers 6 wks. bef. mtg.

LOUISIANA-MISSISSIPPI, Buena Vista Hotel-Motel, Biloxi, Mississippi, February 17–18, 1978.

MARYLAND-DISTRICT OF COLUMBIA-VIRGINIA, Saturday before Thanksgiving and last Saturday in April.

METROPOLITAN NEW YORK, late April or early May 1978. Deadline for papers 2 wks. bef. mtg.

MICHIGAN, Michigan State University, East Lansing, May 5–6, 1978.

MISSOURI, Central Missouri State University, Warrensburg, April 7–8, 1978.

NEBRASKA, University of Nebraska at Omaha, April 14–15, 1978.

NEW JERSEY, Steinhart High School, Trenton, April 28, 1978.

NORTH CENTRAL, College of St. Thomas, St. Paul, Minnesota, April 21–22, 1978.

NORTHEASTERN, Saturday after Thanksgiving, and third week in June in odd-numbered years.

NORTHERN CALIFORNIA, College of Notre Dame, Belmont, February 18, 1978.

OHIO, The University of Akron, April 28–29, 1978.

OKLAHOMA-ARKANSAS, Henderson State University, Arkadelphia, Arkansas, March 31–April 1, 1978.

PACIFIC NORTHWEST, University of Oregon, Eugene, June 16–17, 1978.

PHILADELPHIA, Saturday before Thanksgiving.

ROCKY MOUNTAIN, South Dakota School of Mines and Technology, Rapid City, April 28–29, 1978.

SEAWAY, Brock University, St. Catharines, Ontario, Canada, May 5–6, 1978.

SOUTHEASTERN, Clemson University, Clemson, South Carolina, March 31–April 1, 1978.

SOUTHERN CALIFORNIA, first or second Saturday in March.

SOUTHWESTERN, New Mexico Institute of Mining and Technology, Socorro, Spring 1978.

TEXAS, Stephen F. Austin State University, Nacogdoches, March 31–April 1, 1978.

WISCONSIN, University of Wisconsin-Whitewater, late April 1978.

FUTURE MEETINGS OF OTHER ORGANIZATIONS

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, Washington, February 12–17, 1978.

AMERICAN MATHEMATICAL ASSOCIATION OF TWO YEAR COLLEGES

AMERICAN MATHEMATICAL SOCIETY, Atlanta, Georgia, January 4–7, 1978.

AMERICAN SOCIETY FOR ENGINEERING EDUCATION

ASSOCIATION FOR COMPUTING MACHINERY

ASSOCIATION FOR SYMBOLIC LOGIC, Sheraton-Park, Washington, D.C., December 28–29, 1977.

ASSOCIATION FOR WOMEN IN MATHEMATICS, Atlanta, Georgia, January 4–8, 1978.

CANADIAN SOCIETY FOR HISTORY AND PHILOSOPHY OF MATHEMATICS/SOCIÉTÉ CANADIENNE D'HISTOIRE ET DE PHILOSOPHIE DES MATHÉMATIQUES

FIBONACCI ASSOCIATION

INSTITUTE OF MATHEMATICAL STATISTICS

MU ALPHA THETA

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS, San Diego, California, April 12–15, 1978.

OPERATIONS RESEARCH SOCIETY OF AMERICA, American Hotel, New York City, May 1–3, 1978 (Joint Meeting with the Institute of Management Sciences).

PI MU EPSILON

SCHOOL SCIENCE AND MATHEMATICS ASSOCIATION

SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, University of Wisconsin, Madison, May 24–26, 1978.

THE CARUS MATHEMATICAL MONOGRAPHS

The Monographs are a series of expository books intended to make topics in pure and applied mathematics accessible to teachers and students of mathematics and also to non-specialists and scientific workers in other fields.

These numbers are currently available:

1. *Calculus of Variations,* by G. A. Bliss.
2. *Analytic Functions of a Complex Variable,* by D. R. Curtiss.
3. *Mathematical Statistics,* by H. L. Rietz.
4. *Projective Geometry,* by J. W. Young.
6. *Fourier Series and Orthogonal Polynomials,* by Dunham Jackson.
7. *Vectors and Matrices,* by C. C. MacDuffee.
8. *Rings and Ideals,* by N. H. McCoy.
9. *The Theory of Algebraic Numbers* (Second edition), by Harry Pollard and Harold G. Diamond.
10. *The Arithmetic Theory of Quadratic Forms,* by B. W. Jones.
11. *Irrational Numbers,* by Ivan Niven.
12. *Statistical Independence in Probability, Analysis and Number Theory,* by Mark Kac.
13. *A Primer of Real Functions* (Second edition), by Ralph P. Boas, Jr.
14. *Combinatorial Mathematics,* by H. J. Ryser.
15. *Noncommutative Rings,* by I. N. Herstein.
16. *Dedekind Sums,* by Hans Rademacher and Emil Grosswald.
17. *The Schwarz Function and its Applications,* by Philip J. Davis.
18. *Celestial Mechanics,* by Harry Pollard.

One copy of each Carus Monograph may be purchased by individual members of the Association for $6.50 each; additional copies and copies for nonmembers are priced at $11.00 each. (Orders for under $10.00 must be accompanied by payment. Prepaid orders will be delivered postage and handling free.)

Orders should be sent to:

MATHEMATICAL ASSOCIATION OF AMERICA
1225 Connecticut Avenue, N.W.
Washington, D.C. 20036

NEWS AND NOTICES

PERSONAL ITEMS

GENERAL INFORMATION

NECROLOGY

REPORTS AND ANNOUNCEMENTS OF THE ASSOCIATION AND ITS SECTIONS

MEETINGS AND ANNOUNCEMENTS OF THE ASSOCIATION

MEETINGS OF ITS SECTIONS

CHANGE OF ADDRESS — AMERICAN MATHEMATICAL MONTHLY

OLD ADDRESS

Please affix address label from a recent issue here. The codes it contains are essential for prompt and accurate processing of your requested change.

NEW ADDRESS

NAME

ADDRESS

CITY STATE ZIP

Please return this form to

The Mathematical Association of America
1225 Connecticut Ave., NW
Washington, D.C. 20036

Apply these to your math courses...

ORDINARY DIFFERENTIAL EQUATIONS
Tyn Myint-U, Manhattan College
Covering important methods and techniques, this is an excellent introduction to the theory of ordinary differential equations. Numerous examples are cited, and the exercises range from solving routine problems to proving more difficult theorems. 1977 288 pages $18.50 ISBN 0-444-00233-2

NUMERICAL ANALYSIS FOR COMPUTER SCIENCE
Irving A. Dodes, Kingsborough College, CUNY
Here's the first serious exploration of the mathematical background of computer programming. Intended for a one-semester or one-year course, the book serves as a text and reference for computer science majors. November 1977 640 pages $18.95 ISBN 0-444-00238-3

A COURSE IN MATHEMATICAL LOGIC
J.L. Bell, London School of Economics and **M. Machover,** Chelsea College
A comprehensive, up-to-date text for one-year graduate (or advanced undergraduate) courses in mathematical logic and the foundations of mathematics. Designed for self-study, the text requires no previous knowledge of logic. 1977 600 pages $19.50 ISBN 0-7204-2844-0

INTRODUCTORY COMBINATORICS
Richard A. Brualdi, University of Wisconsin
A stimulating, fresh approach, this is a comprehensive introduction to combinatorics or discrete mathematics. Thoroughly class-tested and containing over 500 student exercises, the book is an ideal text for courses in the ideas and methods of combinatorics.
"... timely, comprehensive, useful, important and welcome."—*Frank Harary, University of Michigan*
1977 384 pages $18.50 ISBN 0-7204-8610-6

AN INTRODUCTION TO LINEAR ALGEBRA FOR SCIENCE AND ENGINEERING, 2nd Ed.
Dominic G.B. Edelen, Lehigh University and **Anastasios D. Kydoniefs,** Lehigh University
The revised and updated edition of this widely adopted text is designed to engage the student in a dialogue. The mathematics is precise, yet the style is informal. Presenting an ideal balance between conceptual fundamentals and computational procedures, the book offers a clear presentation of theoretical principles and their applications. 1976 288 pages $10.95 ISBN 0-444-00195-6

GRAPH THEORY WITH APPLICATIONS
J.A. Bondy, University of Waterloo and **U.S.R. Murty,** University of Waterloo
"A well-written book with an excellent selection of topics."—*L.W. Beineke, Purdue University*
1976 224 pages $19.50 ISBN 0-444-19451-7

MODEL THEORY, 2nd Ed.
C.C. Chang, University of California and **H.J. Keisler,** University of Wisconsin
"... For the first time, a definitive and encyclopedic coverage ... a significant and massive undertaking ..."—*The American Mathematical Monthly*
1977 566 pages Paper $19.50 ISBN 0-7204-0692-7

For more information, or to be considered for complimentary examination copies, write to Cheryl Hagen, Dept. AMM, N.Y. office. Please include your course name, enrollment, and title of present text.

ELSEVIER NORTH-HOLLAND INC
52 VANDERBILT AVENUE NEW YORK, NEW YORK 10017